

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 25, 2004, 04:04:07 ; Search time 7867.05 Seconds

(without alignments)
10792.997 Million cell updates/sec

Title: US-09-893-033-6

Sequence: 1959

Sequence: 1 gccatagagagccatcgccg.....cataagccaatgaaatc 1959

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 2167151695 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: gb_ba:*
2: gb_hcg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pac:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_scs:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_scs:*
28: em_un:*
29: em_vl:*
30: em_hcg_hum:*
31: em_hcg_inu:*
32: em_hcg_other:*
33: em_hcg_mus:*
34: em_hcg_pin:*
35: em_hcg_rtd:*
36: em_hcg_mam:*
37: em_hcg_vrt:*
38: em_sy:*
39: em_higo_hum:*
40: em_higo_mus:*
41: em_higo_other:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1905.2	97.3	146174	1 D90910	D90910 Synchocyst
2	924.6	47.2	1836	1 SPUNASX	M75906 Spirulina p
3	878.4	44.8	300300	1 AP005373	AP005373 Theomycin
4	849.2	43.3	342850	1 AP003597	AP003597 Noctoc sp
5	683.4	34.9	1737	1 SPUNASX	M75907 Spirulina p
6	672	34.3	297800	1 AP006579	AP006579 Gloeobact
7	647.4	33.0	349213	1 BX569693	BX569693 Synchococ
8	638.4	32.6	2338	1 PSP224709	PSP224709 Porphyrid
9	636	32.5	346683	1 BX572098	BX572098 Prochloro
10	603.6	30.8	4099	8 PTECPTRA	M94625 Porphyra um
11	603.6	30.8	191028	8 PTECPTRA	U38804 Porphyra pu
12	592.8	30.3	149987	8 AB002583	AB002583 Cyanidol
13	592.6	30.3	121524	8 AF041468	AF041468 Cyanidol
14	584	29.8	164921	8 AF022186	AF022186 Cyanidol
15	566	28.9	300355	1 AE017162	AE017162 Prochloro
16	562.6	28.7	349082	1 BX572091	BX572091 Prochloro
17	417.6	21.3	300272	1 AE017213	AE017213 Geobacter
18	416.6	21.3	1740	6 AR388434	AR388434 Sequence
19	415	21.2	10029	1 AE013966	AE013966 Yersinia
20	415	21.2	201050	1 AJ414143	AJ414143 Yersinia
21	411.6	21.0	11759	1 AE011091	AE011091 Methanosa
22	410.2	20.9	11318	1 AE012835	AE012835 Chlorobiu
23	392	20.0	10029	1 SC0939124	SC0939124 Streptomy
24	382.8	19.5	308050	1 BX571871	BX571871 Phototrab
25	382.6	19.5	349944	1 AX770909	AX770909 Sequence
26	382.6	19.5	349980	6 AX770909	AX770909 Sequence
27	381.6	19.5	44882	1 MLCB637	M75906 Synchocyst
28	381.6	19.5	348950	1 MLEPRTN7	AE013291 Methanosa
29	381	19.4	5806	1 AF117227	AF117227 Salmoneil
30	379.6	19.4	11367	1 AE004317	AE004317 Vibrio ch
31	377.8	19.3	22348	1 AE008699	AE008699 Salmoneil
32	377.8	19.3	251050	1 AL627265	AL627265 Salmoneil
33	377.8	19.3	300169	1 AE016834	AE016834 Salmoneil
34	377	19.2	12167	1 AE005184	AE005184 Escherich
35	377	19.2	281530	1 AP002550	AP002550 Escherich
36	377	19.2	300409	1 AE016755	AE016755 Escherich
37	373.8	19.1	12065	1 AE004883	AE004883 Pseudomon
38	373.6	19.1	300933	1 AE016791	AE016791 Pseudomon
39	370.2	18.9	4655	1 AE015045	AE015045 Shigella
40	370.2	18.9	290029	1 AE016978	AE016978 Shigella
41	368.6	18.8	1734	6 E11874	E11874 Genomic DNA
42	368.6	18.8	21757	1 AE000118	AE000118 Escherich
43	368.2	18.8	96109	6 AR408756	AR408756 Sequence
44	368.2	18.8	96109	6 AX067460	AX067460 Sequence
45	367.8	18.8	10029	1 AE006125	AE006125 Pasteurel

ALIGNMENTS

RESULT 1
LOCUS D90910 146174 bp DNA linear BCT 04-JUL-2001
DEFINITION Synchocystis sp. PCC 6803 DNA, complete genome, section:12/27,
1430419-1576592.
ACCESSION D90910 AB001339 BA000022
VERSION D90910.1 GI:1652956
KEYWORDS
SOURCE
ORANISM Synchocystis sp. PCC 6803
Bacteria; Cyanobacteria; Chroococcales; Synchocystis.
REFERENCE
AUTHORS Kaneke,T., Tanaka,A., Sato,S., Kotani,H., Sazuka,T., Miyajima,N.,
Sugiura,M. and Tabata,S.
TITLE Sequence analysis of the genome of the unicellular cyanobacterium

FEATURES	COMMENT
JOURNAL	Synechocystis sp. strain PCC6803. I. Sequence features in the 1 Mb
JOURNAL	region from map positions 64% to 92% of the genome
MEDLINE	DNA Res. 2 (4), 153-166 (1995)
PUBMED	96127529
REFERENCE	8590279
AUTHORS	2
TITLE	Kanehisa, T., Sato, S., Kozaki, H., Tanaka, A., Asamizu, E., Nakamura, Y., Miyajima, N., Hirose, M., Sugita, M., Sasamoto, S., Kimura, T., Hosouchi, T., Matsuno, A., Muraki, A., Nakazaki, N., Nambu, K., Okumura, S., Shimpo, S., Takeuchi, C., Wada, T., Watanabe, A., Yamada, M., Yasuda, M. and Tabata, S.
JOURNAL	Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis sp. strain PCC6803. II. Sequence determination of the entire genome and assignment of potential protein-coding regions
MEDLINE	DNA Res. 3 (3), 109-136 (1996)
PUBMED	97061201
REFERENCE	8905231
AUTHORS	3 (bases 1 to 146174)
TITLE	Tabata, S.
JOURNAL	Direct Submission
COMMENT	Submitted (28-JUN-1996) Satoshi Tabata, Kazusa DNA Research Institute, The First Laboratory for Plant Gene Research, Yana 1532-3, Katsarazu, Chiba 292-0812, Japan (E-mail: ktabata@kazusa.or.jp, http://www.kazusa.or.jp/cyano/ , Tel:81-438-52-3933 (ex.2330), Fax:81-438-52-3934) Potential protein coding regions were assigned on the basis of similarity search of the ORFs and Genemark analysis.
FEATURES	Location/Qualifiers
source	1..146174
gene	/organism="Synechocystis sp. PCC 6803"
CDS	/mol_type="genomic DNA"
gene	/strain="PCC6803"
CDS	/db_xref="taxon:1148"
gene	/note="synonym:Synechocystis PCC6803"
CDS	1..486
gene	/gene="apca"
CDS	1..486
gene	/gene="apca"
CDS	/note="ORF ID:slr2067"
gene	/codon_start=1
CDS	/transl_table=1
gene	/product="allophycocyanin a chain"
CDS	/protein_id="BA17874.1"
gene	/db_xref="GI:1652957"
CDS	/translation="MSISIVNADAEARYLSPGELIRIKAFVTGAARLRIAEITLL GSERTIVAGADRLEFORPDIIVSPGNAYEENATCLDMOYLRILVTGVGSDVTV PIREIIGLVGAEVEMRSLGTPIEVAQSVREKVEVAGLMSDDAAEAAYDPVIGKX S"
gene	582..1067
gene	/gene="apcb"
CDS	582..1067
gene	/gene="apcb"
CDS	/note="ORF ID:slr1986"
gene	/codon_start=1
CDS	/transl_table=1
gene	/product="allophycocyanin b chain"
CDS	/protein_id="BA17875.1"
gene	/db_xref="GI:1652958"
CDS	/translation="MDAIVTAVNSADVQGYLDGAAMDKLKSPASGELVRRASAVYI SNATATIKAEAVAKSLIYSDVTRGMMVTRTRVACIDLDYIRATVATMLAGDAS IIDEVNLTKETYNLSIGVPISSIVQAIAIKETVATSLVGADAGKEMGVYDIOSG S"
gene	1280..1483
gene	/gene="apcc"
CDS	1280..1483
gene	/gene="apcc"
CDS	/note="ORF ID:ssr3383"
gene	/codon_start=1
CDS	/transl_table=1
gene	/product="phycobilisome LC linker polypeptide"
CDS	/protein_id="BA17876.1"
gene	/db_xref="GI:1652959"
CDS	/translation="VRRMFRITACVPSQTRIRIORTYONTYFTKLVPYDNNWREDOIRIM

[illegible]

```

/Note="ORF_ID:slr1992"
/codon_start=1
/transl_table=11
/product="glutathione peroxidase"
/protein_id="BA017881.1"
/db_xref="GI:1652964"
/translation="MPLPSLTLLDGTPLAPRVADKIVLPVNVASKGLPQYSGLV
ADKAYGEKAVITIGVPCNFGAEPGSPERIKDTKIKIDVDLTLLERKQVNGNRS
PLVQPLVGDGEIISWMPKFLIGRDGVVAFDPQTKPDDTLKALERKALG"
6955. .8184
/gene="chl"
6955. .8184
/gene="chl"
/Note="ORF_ID:slr1993"
/gene="chl or chl"
/codon_start=1
/transl_table=11
/product="acetyl coenzyme A acetyltransferase (thiolase)"
/protein_id="BA017882.1"
/db_xref="GI:1652965"
/translation="MAAHNPWMHAKTIKRDVPIYAAKTPPLGRPGSLTNSADLGA
HWKSYLAQVGVGDOLDIYIGNVLRAGHGLIPROAKLABIPDIVGFAVDMVCS
SAMSVINALITRAGEBDLIAGETESNSQGYLSHRAMKAEFLMGAEENITDL
LHDGLDSTNGEMGEOTEKLAHEGFSEIIEDEVALCSQORAAHATESGYDSIAP
EITSRKGTQVLASDEGIRSDTTVESLGRSAFAPDGLVLRNCSQITDGAALLLA
SGEAVEKQIKPLAKILGSSMAAGTSPSRPELPITASOKLAKIDKTLADPLPENNE
ARFSVNLPERLGVDRDKLVNNGAIALGHPIGASGARINVTLLYALQQRDKLGLA
ALCHGTGGSTALALERV"
8286. .9008
/gene="fabG"
8286. .9008
/gene="fabG"
/Note="ORF_ID:slr1994"
/codon_start=1
/transl_table=11
/product="3-ketoacyl-acyl carrier protein reductase"
/protein_id="BA017883.1"
/db_xref="GI:1652966"
/translation="MLSLGLSDKVIIVTGNRGIGALIVKLQEWKAVAFETDATTG
GNTBALGVANVTDESMTPAAAEITDKLGPYGVANAGITKDNFPEKLPALMDAV
LNVNLKVAIVSIKPIEBMYERKAGSIYALISISGRNVCQTNSATKAVITGMKS
Query Match 97.3%; Score 1905.2; DB 1; Length 146174;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1943; Conservative 0; Mismatches 9; Indels 3; Gaps 3;

```

```

QY 426 GATTACTGGAAGTGGGCGGTGCATGATTTGTAGCGATGCTTTCAGGAAATTGACAT 485
Db 118808 GATTACTGGAAGTGGGCGGTGCATGATTTGTAGCGATGCTTTCAGGAAATTGACAT 118867
QY 486 TTTTGGCATACCTTACCGATCGTTTAAAGCACTCTCTAATGATGATGATGCGGCGGATAT 545
Db 118868 TTTTGGCATACCTTACCGATCGTTTAAAGCACTCTCTAATGATGATGATGCGGCGGATAT 118927
QY 546 GGCTGCATTGTACTAGGCTTTCATCTTGTAGACCGGTCGTCGCCGGGCGGTTT 605
Db 118928 GGCTGCATTGTACTAGGCTTTCATCTTGTAGACCGGTCGTCGCCGGGCGGTTT 118987
QY 606 GATGATATTCCCAAGGATGTGGCTTAAAGAAATGTGATCAATTCCTCGACCCCG 665
Db 118988 GATGATATTCCCAAGGATGTGGCTTAAAGAAATGTGATCAATTCCTCGACCCCG 119047
QY 666 TGACGTATATCTACCGGCTTATCGCCCAAGTTAAAGTATCCCGACAAATTAATGC 725
Db 119048 TGACGTATATCTACCGGCTTATCGCCCAAGTTAAAGTATCCCGACAAATTAATGC 119107
QY 726 GGCATTGCATTTGTTGAGCAGGCGAGAAATCCCTTGTCTACGTAGGGGAGGCGCAT 785
Db 119108 GGCATTGCATTTGTTGAGCAGGCGAGAAATCCCTTGTCTACGTAGGGGAGGCGCAT 119167
QY 786 CGCCGCCAATGCCCATGCCAGGTGCAGAAATTTGGGGAAGGTTCCAGTTGCGCGTAA 845
Db 119168 CGCCGCCAATGCCCATGCCAGGTGCAGAAATTTGGGGAAGGTTCCAGTTGCGCGTAA 119227
QY 846 AACCAACCTGATGGAAATTTGGGCTTTTGAAGAAACATCCCTTGGTGGTATGTT 905
Db 119228 AACCAACCTGATGGAAATTTGGGCTTTTGAAGAAACATCCCTTGGTGGTATGTT 119287
QY 906 GGGATGATGATGCGACCGCTATGCACTTGTGCGTACAGGATGATGATTTGTTGATTC 965
Db 119288 GGGATGATGATGCGACCGCTATGCACTTGTGCGTACAGGATGATGATTTGTTGATTC 119347
QY 966 AGTGGGGCGCGTTTTCAGACACCGGGGTAATGCGAAACTAGAGAAATTTGCTAGCGCG 1025
Db 119348 AGTGGGGCGCGTTTTCAGACACCGGGGTAATGCGAAACTAGAGAAATTTGCTAGCGCG 119407
QY 1026 CAAAGTATTCATTCATGACATCGACCGGCGGAGGTGGAAAAACAGGCTCCCATGT 1085
Db 119408 CAAAGTATTCATTCATGACATCGACCGGCGGAGGTGGAAAAACAGGCTCCCATGT 119467
QY 1086 GCCCATGTGGGGGATGTACGCAATGTTTAAAGACGTTTTCAGAGCGGCGCGGAAT 1145
Db 119468 GCCCATGTGGGGGATGTACGCAATGTTTAAAGACGTTTTCAGAGCGGCGCGGAAT 119527
QY 1146 GGATTTACCCACCCATCCCATACCAACCAAGGATGTTAAATGCAATTGATGATGCG 1205
Db 119528 GGATTTACCCACCCATCCCATACCAACCAAGGATGTTAAATGCAATTGATGATGCG 119587
QY 1206 GACCGATTTACCCCTCCAGGTGCCCATCTATGAGAACTATTTGCCCCCAAGAGATGT 1265
Db 119588 GACCGATTTACCCCTCCAGGTGCCCATCTATGAGAACTATTTGCCCCCAAGAGATGT 119647
QY 1266 ACAGAAATTTGTGCGCAGGCGCCCGCATGCTCTACACCAACGATGTTGGGAACAACCA 1325
Db 119648 ACAGAAATTTGTGCGCAGGCGCCCGCATGCTCTACACCAACGATGTTGGGAACAACCA 119707
QY 1326 AATGTGGCGCGCCAGTTTGTAAACAATGGCCCCCGCATGATGATTTTCAGTGTGCTT 1385
Db 119708 AATGTGGCGCGCCAGTTTGTAAACAATGGCCCCCGCATGATGATTTTCAGTGTGCTT 119767
QY 1386 GGGTACATGAGGCTTGTGTTTACCTGCGCATGAGGACCAAAAGTGGGATGAGGAGCG- 1444
Db 119768 GGGTACATGAGGCTTGTGTTTACCTGCGCATGAGGACCAAAAGTGGGATGAGGAGCGA 119827
QY 1445 AGCGGTATTTGCATCAGTGGAGATGCGACCTTCAAAATGAATCTTCAGGAATCGGAGAC 1504
Db 119828 AGCGGTATTTGCATCAGTGGAGATGCGACCTTCAAAATGAATCTTCAGGAATCGGAGAC 119887

```

QY 1505 CCTAGCCAGTACGATCCAGTAACTATTCTCATTAAGCGTTGGCAGGGAT 1564
 Db 119888 CCTAGCCAGTACGATCCAGTAACTATTCTCATTAAGCGTTGGCAGGGAT 119947
 QY 1565 GGTGCGTCACTGCGCAACAACTTTCTACGAAAGCGTTATTCGTTCTAACTGTC 1624
 Db 119948 GGTGCGTCACTGCGCAACAACTTTCTACGAAAGCGTTATTCGTTCTAACTGTC 120007
 QY 1625 GGGCATGCGCAACATTAATCTCTCTGGAAGCGTATGATGATGATGATG 1684
 Db 120008 GGGCATGCGCAACATTAATCTCTCTGGAAGCGTATGATGATGATGATG 120067
 QY 1685 CAAGCGGGAAGATTGGCCCGCGCGATGCGGAAATGCTAGCCCAATGGTCTGT 1744
 Db 120068 CAAGCGGGAAGATTGGCCCGCGCGATGCGGAAATGCTAGCCCAATGGTCTGT 120127
 QY 1745 GATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1804
 Db 120128 GATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120187
 QY 1805 TAATGCCCAATGCTAGCTTACCGGAGTGCCTGTCAGGACATGGTCCCGGATGT 1864
 Db 120188 TAATGCCCAATGCTAGCTTACCGGAGTGCCTGTCAGGACATGGTCCCGGATGT 120247
 QY 1865 GAGAGTGCACATTCCTCCCAACCCCAATTTTATCATCCCATGCTTCTGTTGTGG 1924
 Db 120248 GAGAGTGCACATTCCTCCCAACCCCAATTTTATCATCCCATGCTTCTGTTGTGG 120307
 QY 1925 AGCCAACTCTTACCCCATTAAGCCAAATTTGAATTC 1959
 Db 120308 AGCCAACTCTTACCCCATTAAGCCAAATTTGAATTC 120342

RESULT 2
 SPUNHASK
 DEFINITION 1836 bp DNA linear BCT 26-Apr-1993
 ACCESSION M75906
 VERSION M75906.1 GI:152904
 KEYWORDS acetylhydroxy acid synthetase.
 SOURCE Spirulina platensis
 ORGANISM Bacteria; Cyanobacteria; Oscillatoriales; Spirulina.
 REFERENCE 1 (bases 1 to 1836)
 AUTHORS Milano, A., De Rossi, E., Zanaria, E., Barbierato, L., Ciferri, O. and Ricciardi, G.
 TITLE Molecular characterization of the genes encoding acetylhydroxy acid synthase in the cyanobacterium *Spirulina platensis*
 JOURNAL J. Gen. Microbiol. 138 (Pt 7), 1399-1408 (1992)
 MEDLINE 92381487
 PUBMED 1512571
 COMMENT Original source text: *Spirulina platensis* (strain C1) DNA.
 FEATURES
 source location/Qualifiers
 1..1836
 /organism="Spirulina platensis"
 /mol_type="genomic DNA"
 /strain="C1"
 /db_xref="taxon:1156"
 2..1836
 /gene="11vX"
 <1..>1836
 /gene="11vX"
 /EC_number="4.1.3.18"
 /codon_start=1
 /transl_table=11
 /product="acetylhydroxy acid synthase (AHAS)"
 /protein_id="AAA26594.1"
 /db_xref="GI:152905"
 /translation="VOLOTKIANKARATGAFELIDSLKRGVQIFGVPGALIPIDE
 LYRAEAEIOHILVRHOGASHADGVARATGAVGCFGSGGADNLTGATAM
 DSIPIVITGQVAPALGTADSDIFGLTPIVKASYVRBEGDANLVAEFHIA
 STGRPGVLIIVPDVGEEDFYI PVNGEVSIPGYRPIYKANROINOKLIEBE
 RPLMYVGGATISATNAHAIADLLELFOIPYTTLMGSGSDEKRPISVGLMGHGTAY

ORIGIN

Query Match 47.2%; Score 924.6; DB 1; Length 1836;
 Best Local Similarity 69.7%; Pred. No. 1.9e-259;
 Matches 1265; Conservative 0; Mismatches 550; Indels 1; Gaps 1;

QY 119 TTCTTAGCCACGGAACCGGGCTTATCTCTGATGATGATGATGATGATGATG 178
 Db 20 TTGCTGCGCAAGCGTGCACAGCTGGCTTTCGCTTATGATGATGATGATGATG 79
 QY 179 TCAACACATTTTGGGCTATCCCGCGGGGCAATTTGCGCATATATGATGATGATG 238
 Db 80 TACAGCACATTTTGGCTATCTGCTGATGATGATGATGATGATGATGATGATG 139
 QY 239 GCTTTGAGCGGCGGGGGAATTTGAGCATTTTGGTGGCGCCATGACAGAGGCTTCC 298
 Db 140 GAGCGGAAGCCGAGGCGCATATACAGCATATCTGATGATGATGATGATGATG 199
 QY 299 ATGCGCGGATGGGATATGCGAGGCGACAGTAAAGTGGAGTTGTTGGTATATG 358
 Db 200 AGCGCGGATGGGATATGCGAGGCGACAGTAAAGTGGAGTTGTTGGTATATG 259
 QY 359 GACCAAGGCGCATCTGATGATGATGATGATGATGATGATGATGATGATGATG 418
 Db 260 GACCAAGGCGCATCTGATGATGATGATGATGATGATGATGATGATGATGATG 319
 QY 419 TGGTGGTATTTATGAGAGTGGGCGTGGCATATTTGATGATGATGATGATGATG 478
 Db 320 TGGTGGTATTTATGAGAGTGGGCGTGGCATATTTGATGATGATGATGATGATG 379
 QY 479 TTGACATTTTGGCATATCTTACCATGATGATGATGATGATGATGATGATGATG 538
 Db 380 CAGATATTTTGGCATATCTTACCATGATGATGATGATGATGATGATGATGATG 439
 QY 539 CGGATATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 598
 Db 440 GTGACATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 499
 QY 599 CGGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 658
 Db 500 CGGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 559
 QY 659 ACCCGGTGACGTTATCTTACCGGCTTATCGCCCAAGGTTAAAGTATCCCGCA 718
 Db 560 ATCCAGAGAGGTTCTTACCTGCTATCGCCCAAGGTTAAAGTATCCCGCA 619
 QY 719 TTAAGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 778
 Db 620 TCAACCAAGCATTAATCTTATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 679
 QY 779 GGGCGATCGCGCCCAATGCGCCAGGTCAGAGATTTGGGAAAGTTCCAGTTGC 838
 Db 680 GGGCGATTTCCGCGCCAGGTCAGAGATTTGGGAAAGTTCCAGTTGC 739
 QY 839 CGGTAACCAACCTGATGAGAAATTTGGGCTTTTGAAGAAACCATCCCTTTCG 898
 Db 740 CGGTAACCAACCTGATGAGAAATTTGGGCTTTTGAAGAAACCATCCCTTTCG 799
 QY 899 GATGTTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 958
 Db 800 GATGTTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 859
 QY 959 TGAATGACATGAGGCGCGTTTGAAGAGCGGGTAACTGGCAATCAAGATTTGCT 1018
 Db 860 TGAATGACATGAGGCGCGTTTGAAGAGCGGGTAACTGGCAATCAAGATTTGCT 919
 QY 1019 GCGCGCGCAAGATTAATCAATTAATGATGATGATGATGATGATGATGATGATG 1078

Db 920 CTAGGGCGAAGGTGATTCATATTTGACATTTGACCCCGGAACTTGTAAACCCGAGCC 979

Qy 1079 CCGATGTGCGCATTTGTGGGGAGATGATACCGCATTTTGTGAACAGCTTTTGCAGGGGGCC 1138

Db 980 CCGAAGTGCATTTGTGCGAGATGTTCGCGAGGTGTATACACTGTGGCGCGCTGCC 1039

Qy 1139 GGGAAATGTGATTACCCCAACCCATCCCATCCAGCGATGTGTTAAATGECATTATC 1198

Db 1040 GAGAAATGTGAGATGTGCGGGAACGATTAATCAAACTCAATCGGTGTTAGCGGATTAC 1099

Qy 1199 ATTGGCGAGCGATTAACCCCTCCAGTGGCCCACTATGAGAGATCTATGCCCCCAG 1258

Db 1100 GTTGGCCGGAAGATTATCTCTGTGGTTCCTCTTATTTGATGACCTCTCTCCCAAG 1159

Qy 1259 AGGTAGTACACGAAATTTGTCGCCAGGCCCGCATGCTACTACACCAACCGATGTGGAC 1318

Db 1160 AGGTATGCGCAGAAATTAGGAAAAATGGCAGCGGATGGTACTATACACGAGATGGGTC 1219

Qy 1319 AACACCAATTTGGGGCGGCCAGTTTGTGAACAATGGCCCCCGCGATGATTTCCAGTG 1378

Db 1220 AACACCAATTTGGGGCGGCTCAATTTTGAATAATGGCCCCCGTCAATGATATCTAC 1279

Qy 1379 CTGGCTTGGTACGATTTGGCTTTGTTTACTGCGCCCATGGAGCCAAAGTGGAGTGG 1438

Db 1280 CTGGCTTGGAGTATGAGCTATGAGCTATGATGCTGCGGCTATGAGAGCTAAAGTGGCTTGG 1339

Qy 1439 -GGAGCAGCGGTCAATTTGATCAATGAGATGCGAGCTTCCAAATGAATCTTGAGAAC 1497

Db 1340 AGCGATCGCAGGTGATTTGTATGCTGGGGATGCCAGTGTCAATGAATCCAGAAC 1399

Qy 1498 TGGGAACCCCTAGCCCAAGTACGACATCCAGTTAAACTATTTCTCAATAACGTTGGC 1557

Db 1400 TGGGAACATACGACATTTGGCAATTAATGTAACACGTAATTAATTAACGTTGGC 1459

Qy 1558 AGGGATGTGCTGATGATGCGCAACAACTTTCTACGAAAGCTATTTCTGCTTAAACA 1617

Db 1460 AAGGATGTGCTGCGCAGTGGCAACAGCTTTTATGAGAGCTTATTCGGTTCTAACA 1519

Qy 1618 TGTCCAGGGCATGCGCAACATTAATCTCTCTGTGAAAGCTATGCGATCAAGGTTATTA 1677

Db 1520 TGGAAATCGGAATGCGCAATTTTGAATGTTAGCACTTCCATGCGCTAAAGGCATGG 1579

Qy 1678 CTGTGCGCAAGCGGAAATTTGGCCCGCGCATGCGCAAAATGCTAGCCCAATGTGTC 1737

Db 1580 TCGTTAAGATGCTGATGACGATTAACACCAAGCATGACGAAATGTTGCTTATGATGCTC 1639

Qy 1738 CTGTGCTGATGATGATGCTGCTCAAAAAAAGATGAAACTGTTACCTATGATGCCCCG 1797

Db 1640 CTGTGCTGATGATGATGCTGCTCAAAAAAAGATGAAACTGTTATCTATGATGACACAG 1699

Qy 1798 GCATGATGATGATGATGCTGCTGCTCAAAAAAAGATGAAACTGTTACCTATGATGCCCC 1857

Db 1700 GAGAGATGATGCTGATGATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1759

Qy 1858 GAGATGATGATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1917

Db 1760 AGTTAATCTATGATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1819

Qy 1918 GTTGTGAGAGCAAACT 1933

Db 1820 AGTGTGACCTAACT 1835

RESULT 3
AP005373 300300 bp DNA linear BCT 17-AUG-2002
LOCUS Thermosynechococcus elongatus BP-1 DNA, complete genome, section
DEFINITION 5/9
ACCESSION AP005373 BA000039
VERSION AP005373.1 GI:22294885
KEYWORDS
SOURCE Thermosynechococcus elongatus BP-1

ORGANISM
Thermosynechococcus elongatus BP-1
REFERENCE
Bacteria; Cyanobacteria; Chroococcales; Thermosynechococcus.
AUTHORS
1 Nakamura, Y., Kaneko, T., Sato, S., Ikeuchi, M., Katoh, H., Sasaoka, S.,
Kawanabe, A., Iriuguchi, M., Kawashima, K., Kimura, T., Kishida, Y.,
Kiyokawa, C., Kohara, M., Matsumoto, M., Matsuno, A., Nakazaki, N.,
Shimpo, S., Sugimoto, M., Takeuchi, C., Yamada, M., and Tabata, S.
TITLE
Complete genome structure of the thermophilic cyanobacterium
Thermosynechococcus elongatus BP-1
JOURNAL
DNA Res. (2002) In press
REFERENCE
2 (bases 1 to 300300)
AUTHORS
Kaneko, T.
TITLE
Direct Submission
JOURNAL
Submitted (05-JUN-2002) Takakazu Kaneko, Kazusa DNA Research
Institute, The First Laboratory for Plant Gene Research, 2-6-7
Kazusa-kametarai, Kisarazu, Chiba 292-0812, Japan
(E-mail: kaneko@kazusa.or.jp,
URL: http://www.kazusa.or.jp/cyano/Thermo/
Tel: 81-438-52-3935 (ex. 2338), Fax: 81-438-52-3934)
FEATURES
source
1. 300300
/organism="Thermosynechococcus elongatus BP-1"
/mol_type="genomic DNA"
/strain="BP-1"
/db_xref="taxon:197221"
/note="BP-1"
complement (293. 1519)
/gene="t111162"
complement (293. 1519)
/gene="t111162"
/note="ORF ID: t111162
putative transposase"
/codon_start=1
/transl_table=11
/protein_id="BAC08714.1"
/db_xref="GI:22294886"
/translation="MTAKRSCVTHRIIAIMSGVEKAFSYRFPPTREOSILRRITGCV
RIYVNRALARTAEWYRKERLDIYOTLSLITQMKKODLOFINEVSCVPLQALRHL
QSNFSPFRGRATVPFKKRNNGSAEFTKSAFRMDKGYFLAKNCPINPMSRRLP
DVEPSTVTIRLPAQWYISLRFDPREITLQVPSVGLDVISLITLTSGEKTA
NPGHFNRYRKLKRAQKRSLSRKQGSRRMDKALVAKIHQKISDKRKHQJLTTTL
IRENQIVIESLAVKIMVNRQJARSISAGWGEIVROLEYKQAWGRTIVKIDRWFP
SSRCGCGHIVEMLPISVREMDPCPGCAHHDHINAAGNIIAVGHTVPCGAVPDP
RHTSGGLRRSRSSQK"
1689. 2072
/gene="t111163"
/note="t111163"
1689. 2072
/gene="t111163"
/note="ORF ID: t11163
hypothetical protein"
/codon_start=1
/transl_table=11
/protein_id="BAC08715.1"
/db_xref="GI:22294887"
/translation="MANTPSRPPPLTVESPRSATAFVHLDSPVMTGLRRSLFIADGT
AVENGOLHTMIALLELYALASKRNQIIVIOEHPSQAPLPFIEVSRORFPVDTPT
LNANLATVLRLEPELFRGSLR"
complement (2082. 2774)
/gene="t111164"
complement (2082. 2774)
/gene="t111164"
/note="ORF ID: t111164
unknown protein"
/codon_start=1
/transl_table=11
/protein_id="BAC08716.1"
/db_xref="GI:22294888"
/translation="MTISPLTRHYVRLLAHPAIOISCIGCDRSRPHKINBEDIDRYIN
FOHUTYPOHQLEFLINLHQKIKANEPHQTLDQHTGRIFDLIGLDLPVSPNTTEL
LSFSEIANDPNKPAAPANOEATLTLPYVELEFRRTIEVAIVLQMGRTIVTNWLLTRP
RESWIAKIQVACFNHDPFPCQQLTNHEAETLQCHINHFIRHCCVLPILRLILLEA
GIPSYLYDLDVD"
3355. 3661
gene

gene
3816..4454
/gene="clt1166"
3816..4454
/gene="clt1166"
/note="ORF ID:clt1166"
/codon_start=1
/transl_table=11
/product="phosphopantetheine adenylyltransferase"
/translation="MLAYPGSFDPTITGHLDIRGARLISEYIVAIANPQKALF
SVGGQIKQQAQATSHLRKNRVDPFDGLTVEARSGRAVLLLRGRVLSDEPEYLQMSH
TKKSLMPELIEYFVLTITSNEYSLVKEIARFGENVRHDPANVAKDEACFTQTP
IPSORPLE"
3816..4454
/gene="clt1166"
3816..4454
/gene="clt1166"
/note="ORF ID:clt1166"
/hypothetical protein"
/codon_start=1
/transl_table=11
/protein_id="BAC08718.1"
/db_xref="gi:22294890"
/translation="MLYSNTPDISITTAGVNEALDAPSAVILQLOQLKEELLILEGT
KIPLGRKILIDEEQILQITQIEEIPBSVEAAQILKQDEIIKKQYKQELILRAA
EORAAIDTELIRQDAIEAQIKIQOOVEIEAQQALRIEIMLRNTEKHEALHR
OYTHSECORQOEDAAVADRTLAEMERQLKEMIAVIQGRQYVLKQHPAHRP"
complement(4548..5525)
/gene="clt1167"
complement(4548..5525)
/gene="clt1167"
complement(4548..5525)
/gene="clt1167"
/note="ORF ID:clt1167"
hypothetical protein"
/codon_start=1
/transl_table=11
/protein_id="BAC08719.1"
/db_xref="gi:22294891"
/translation="MNACQYMLSRVADAVMLNRYIERAENIAPVDVNLMLDLPT
SISQMDPEVLITGDLSPFOEYGTAEENYIAPLIPITYPNSIISGLRARNNAHS
IIEVSIEMQOVNAFYTMVEBAAPKGLIETASPLQEQASHLFGVMQSTMSHNS
AHHQVGMGLIRADKTRIRLDVXYLLIPSEVDGPTIDELGMIALKSAAYEMVR
KGRSRIRPAGVAFELIIDGERRAIRCLLOVEKSIYRITGPTLGSWHQVSEQLQR
LRSQDLIRITIDETISQMHREPLDIYQCOMNEVDIOTFPTLBPVKTR"
5565..5897
/gene="ycf20"
5565..5897
/gene="ycf20"
/note="ORF ID:clt1168"
hypothetical protein"
/codon_start=1
/transl_table=11
/protein_id="BAC08720.1"
/db_xref="gi:22294892"
/translation="MORTRLDLRLCAGIEOOLKNPWRLATLSTIFLEGVLGLIA
ISSAGQGLYDIIASSWVAIAEYISALPISDRKIKRLQTLFGMLNALKRGLLYGF
IYAPLLGS"
6319..7347
/gene="clt1169"
6319..7347
/gene="clt1169"
/note="ORF ID:clt1169"
/codon_start=1
/transl_table=11
/product="pyruvate dehydrogenase E1 component, alpha
subunit"
/protein_id="BAC08721.1"
/db_xref="gi:22294893"
/translation="MVOERSLPSLSIPOTAIRBQGLMYEDWMLGRTFEDKCEMYTY
BORMGFYILNIGQAVSGVYKAMRPDQVYCTRDVHALISGIVAREMALPQR
ATGCKSGRGSHLISAKNPIFGSPAPAEIGPATGAAPOTIKROROMBAKADQVLT
ACPFEGGSSNRQPFECNLNAAIKPLILFVVENNNKMAIGAAHERASSETITTKAKVY
RMVEBVEGDMDIVARIVAEALARPAGGPTLIEALTYRFQSHSLADEDELRSEK
EKEFMIKQDPIIKRGAYLVEOELATGEDIRAIIEQVQAIVEDAVTFPEOSEPSEBEI

[illegible]

Db	147077	GRAM-negative bacterium	NCBI			AP003597.1	Genomic DNA	147077
Db	147137	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	147137
Db	147197	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	147197
Db	147257	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	147257
Db	147317	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	147317
Db	147377	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	147377
Db	147437	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	147437
Db	147497	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	147497
Db	147557	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	147557
Db	147617	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	147617
Db	147677	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	147677
Db	147737	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	147737
Db	147797	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	147797
Db	147857	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	147857
Db	147917	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	147917
Db	147977	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	147977
Db	148037	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	148037
Db	148097	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	148097
Db	148157	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	148157
Db	148217	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	148217
Db	148277	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	148277
Db	148337	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	148337
Db	148397	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	148397
Db	148457	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	148457
Db	148517	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	148517
Db	148577	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	148577
Db	148637	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	148637
Db	148697	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	148697
Db	148757	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	148757
Db	148817	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	148817
Db	148877	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	148877
Db	148937	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	148937
Db	148997	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	148997
Db	149057	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	149057
Db	149117	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	149117
Db	149177	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	149177
Db	149237	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	149237
Db	149297	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	149297
Db	149357	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	149357
Db	149417	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	149417
Db	149477	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	149477
Db	149537	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	149537
Db	149597	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	149597
Db	149657	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	149657
Db	149717	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	149717
Db	149777	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	149777
Db	149837	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	149837
Db	149897	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	149897
Db	149957	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	149957
Db	150017	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	150017
Db	150077	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	150077
Db	150137	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	150137
Db	150197	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	150197
Db	150257	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	150257
Db	150317	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	150317
Db	150377	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	150377
Db	150437	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	150437
Db	150497	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	150497
Db	150557	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	150557
Db	150617	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	150617
Db	150677	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	150677
Db	150737	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	150737
Db	150797	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	150797
Db	150857	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	150857
Db	150917	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	150917
Db	150977	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	150977
Db	151037	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	151037
Db	151097	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	151097
Db	151157	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	151157
Db	151217	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	151217
Db	151277	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	151277
Db	151337	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	151337
Db	151397	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	151397
Db	151457	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	151457
Db	151517	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	151517
Db	151577	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	151577
Db	151637	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	151637
Db	151697	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	151697
Db	151757	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	151757
Db	151817	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	151817
Db	151877	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	151877
Db	151937	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	151937
Db	151997	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	151997
Db	152057	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	152057
Db	152117	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	152117
Db	152177	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	152177
Db	152237	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	152237
Db	152297	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	152297
Db	152357	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	152357
Db	152417	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	152417
Db	152477	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	152477
Db	152537	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	152537
Db	152597	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	152597
Db	152657	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	152657
Db	152717	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	152717
Db	152777	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	152777
Db	152837	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	152837
Db	152897	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	152897
Db	152957	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	152957
Db	153017	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	153017
Db	153077	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	153077
Db	153137	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	153137
Db	153197	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	153197
Db	153257	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	153257
Db	153317	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	153317
Db	153377	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	153377
Db	153437	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	153437
Db	153497	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	153497
Db	153557	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	153557
Db	153617	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	153617
Db	153677	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	153677
Db	153737	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	153737
Db	153797	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	153797
Db	153857	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	153857
Db	153917	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	153917
Db	153977	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	153977
Db	154037	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	154037
Db	154097	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	154097
Db	154157	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	154157
Db	154217	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	154217
Db	154277	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	154277
Db	154337	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	154337
Db	154397	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	154397
Db	154457	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	154457
Db	154517	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	154517
Db	154577	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	154577
Db	154637	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	154637
Db	154697	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	154697
Db	154757	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	154757
Db	154817	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	154817
Db	154877	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	154877
Db	154937	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	154937
Db	154997	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	154997
Db	155057	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	155057
Db	155117	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	155117
Db	155177	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	155177
Db	155237	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	155237
Db	155297	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	155297
Db	155357	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	155357
Db	155417	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	155417
Db	155477	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	155477
Db	155537	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	155537
Db	155597	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	155597
Db	155657	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	155657
Db	155717	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	155717
Db	155777	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	155777
Db	155837	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	155837
Db	155897	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	155897
Db	155957	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	155957
Db	156017	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	156017
Db	156077	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	156077
Db	156137	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	156137
Db	156197	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	156197
Db	156257	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	156257
Db	156317	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	156317
Db	156377	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	156377
Db	156437	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	156437
Db	156497	Gram-negative bacterium	NCBI			AP003597.1	Genomic DNA	156497
Db	156557	Gram-negative bacterium	NCBI			AP003597.1		

```

/codon_start=1
/transl_table=1
/protein_id="BAB76292.1"
/db_xref="GI:17133730"
/translation="MDSKSIQLVDLPSNDITVYNALDYVAPGQWNLVGFNTR
AVGTETARVIOKTRDRASTLYNDPOGQOSAVKLTOTIDKADTAMTALAKVCEK
IGFSLPSNTPKADYVOSTIDLVSKIAETIYVPEKANGITPOPNQEPANSLANNQDA
SLIKFVALVCLDGLPLGPDFLSKIQGISGAESNSTIQNPVLAVNSLPSGSPIDK
LGFINGFNNAVQGMNMLVTKTGITPQSISSHGNFIQIADNIDPVAALDQTPYV
EHTGISOVARLLIKATLVKEIKOETIOPIDSSSAVSDTSQVAVSKIDVWMBE
BEDWQATIEKVDODDFINYIGYSSHNEVWGDIDRADYASDANGAVAGKATVC
MSDEETWYATIEKIQNGYFIHYGTVDSSYDEWVNSDDIS"
1680. 1680
/gene="asr4594"
1680. 1680
/gene="asr4594"
/note="ORF_ID:asr4594
hypothetical protein"
/codon_start=1
/transl_table=1
/protein_id="BAB76293.1"
/db_xref="GI:17133731"
/translation="MTALAPQEKSPQVLLYLGEDNYFVVEVPSLPQCISQKTRBEA
LTNIEVLQDKGELTPEYOFINI"
1877. 2270
/gene="alr4595"
1877. 2270
/gene="alr4595"
/note="ORF_ID:alr4595
unknown protein"
/codon_start=1
/transl_table=1
/protein_id="BAB76294.1"
/db_xref="GI:17133732"
/translation="MSKVDVRLRLFCTVYLGFTTTHVQASAPFYEFINVDIA
VQPNQDMVTEQKNTFRDNLHRSRYTFLDNDVRITVESNGQLSTISTG
NHQLLSMEHQLSPESHTEVLKRVIGLLVNNVAYWEALFSDKAPIDQAVR
VEPEBELTNKIKIQSGVADIRVDTKTIEATXQPIEPGGLKIQVIFDAGTKI
ETSGMSTQFFYERLWVILGLVFGIPLLIIVSRNGTSSSYIGDSSIDYGGGSG
DYGGDYGGDYGGGGGGGG"
3082. 4473
/gene="alr4596"
3082. 4473
/gene="alr4596"
/note="ORF_ID:alr4596"
/codon_start=1
/transl_table=1
/product="malic enzyme"
/protein_id="BAB76295.1"
/db_xref="GI:17133733"
/translation="MADLTNPSFSVTLRQLIPNRVGLASTIQALASGGNIGQIDL
IEQROYSLRDLYVDASTERAETIYQAVKAIPIKLVSYDRTFNLHRSGLISIR
IPLKSVSLMAAYTGVGAVCTAIAQDPOVNLTIKONTVAIVTDSAVLGGNIGP
HALPVMGKMLPKEFAGLDAPICLATQDTEIIOAVKNIAVFGVNLBJIABR
CFTEKRLRDLIDVPFHDDOHTRAIVTLAALNSLKNOKSLADIRIVNGGAGV
AIALILKAKAEKIMCDSKGIISTRTDLNEBQBPAYAOCTLAGAVAGAVPIGV
SAGVLPBMYKSAKDAIVFAMNPIPEIOPELISKVAVIANGSDPNOINNVAA
FPVFGRLDCAQTITTYLXNASAIASLVNPSDLRHHIIPSVDERVAVTAAVA
VQRAAREGIARG"
5270. 6853
/gene="alr4597"
5270. 6853
/gene="alr4597"
/note="ORF_ID:alr4597"
/codon_start=1
/transl_table=1
/product="type I restriction-modification system DNA
methylase"
/protein_id="BAB76296.1"
/db_xref="GI:17133734"
/translation="MTEKAAKQNNSSPEBEKLMKADKLRKNTDASVYKHIVTGLIFL
KYISDAPEELHQKLAGEBDYAGANPEDRDEYSAENVFVPEARWSYLGQCKKQPD
IKTYDLAMEAIEDENAKRLKGLIPKVVGGQCKLDQKSLGGLIDIGSTITGDARAQAD

```

```

VIGREYVFLQGFALAEGRKGQFTYTPESLYKLVLEMLPEYNGRFPDPCGGSGNFVQ
SEKRVKNQGRLDIDISTIGQESNETTYLTCMNLAIRIGIDDSNINKMPEGSFLNDAK
DLKADFLVIANPEFDSWDGSELRLNDGKMLDKDLPVPVGNANFAMVSHPIYHALPTGS
AGFVLSNGSLSSNTSGEDIRKALVQKDLVDCIWMPLQFNTNGIPACLMFLSRYN
GNKNDRGGEVFLIDASELYGVNVRSSAFTEAETISKIADTYHMKSGSGSYRIDKGF
CKSASIAIEKHNPFYLTGRYVGIPEVDEGVSEPEKMSLTALAGQMRGQGLDDEE
IKKQKLKGFIVSDHVDNF"
6838. 7121
/gene="asr4598"
6838. 7121
/gene="asr4598"
/note="ORF_ID:asr4598
unknown protein"
/codon_start=1
/transl_table=1
/protein_id="BAB76297.1"
/db_xref="GI:17133735"
/translation="MMMTQVILKLNLPVIEKLNHLAQSHQRTLEBEITSLLEDVTEN
TPITRSKRDMSGPFPEQTCAGQCELLVRBPQEAQRBDPL"
7226. 7804
/gene="alr4599"
7226. 7804
/gene="alr4599"
/note="ORF_ID:alr4599
hypothetical protein"
/codon_start=1
/transl_table=1
/protein_id="BAB76298.1"
/db_xref="GI:17133736"
/translation="VIAYKGFPKLTPPEYFIWEKQOLRREYINGEVAMTGTQNH
GRIASNIIFIVKGLRGGCGQVNSDRINLFEKDVYDPVSVTCBERDSTAQAO
YPCLIVEVLSPTASYDRGDKFLRYRRNPSIQDYVLVDAEKIAIDLVRKMRGNWEIF
NYOLDNIXKQSIDISFPISSEYDIVEESV"
7855. 8145
/gene="asr4600"
7855. 8145
/gene="asr4600"
/note="ORF_ID:asr4600
unknown protein"
/codon_start=1
/transl_table=1
/protein_id="BAB76299.1"
/db_xref="GI:17133737"
/translation="MIFKEMMTQVILKQNLPIVLEKIKRIAQSHQRTLEBEITSLIE
DVAEKEEPEKSSBGFMDMTLQFRERMOENITFDADPADIIRDSVGRDVEF"
8211. 8573
/gene="alr4601"
8211. 8573
/gene="alr4601"
/note="ORF_ID:alr4601
hypothetical protein"
/codon_start=1
/transl_table=1
/protein_id="BAB76300.1"
/db_xref="GI:17133738"
/translation="MWKLTQHOETATGSMYMEELFGCLRLVESQKRLLEVINQ
IPLKNTILNIDLKQMRHQAERARLSKMGKTPAFIDQIASIANSNLIYLTNNVSDF
ESFNDLAVENFVNSGEG"
8577. 9749
/gene="alr4602"
8577. 9749
/gene="alr4602"
/note="ORF_ID:alr4602"
/codon_start=1
/transl_table=1
/product="type I site-specific deoxyribonuclease chain S"
/protein_id="BAB76301.1"
/db_xref="GI:17133739"
/translation="MSEKKEKTLGEIADIIMWGSPGTGETCNNNGGLPLNGPTEFPG
RNPUPQOTIIPKXIAEGDLFCVRSSTGRMMADOKAYIGGSIASIRAKDILFD
PYRATIRKEIKSLAVATGTFPNISRDHLNLIYQVPSKNITITYSNARITDEKI
YNSRQNETLALAQTLFKHWFIDPEPNADGKPKSSGAMVNSALGYIPEASVCK
LGOYINIKHGYAFRGEYITTEVTEKILTLTPVNFKIGGGFNDSSKXYYSADYSNEVYL

```

RRKDLAIVMTLSKSGBSLGPATIPDIKGVFIHNOIGIKVENNIDKTFYELLR
 REYRSHLIGTSSGSTVHTSPSRICEYFVLPDELIDKFSALATATIDKILLSTL"
 /gene="1.0465"
 /gene="1.0465"
 /gene="1.0465"
 /note="ORF ID: alx4603"
 unknown protein"
 /codon_start=1
 /transl_table=11
 /protein_id="BAB76302.1"
 /db_xref="GI:17133740"
 /translation="MIGLRELINDERNSFSHIALESQAEKXNLLISKVIDLEFV
 KLESISLIQYKXNLSNITDRIPLKFDQSHLSKRNHLLIDNNIRININDNMYRLF

Query Match 43.3%; Score 849.2; DB 1; Length 342850;
 Best Local Similarity 67.0%; Pred. No. 5.9e-237;
 Matches 1235; Conservative 0; Mismatches 604; Indels 4; Gaps 2;

QY	117	TGTTCTAGCCACGGCAACCGGGCTTATATCCTGATGATGCTGAAACGCCATGG	176
DB	24930	TGTGACACCAACACGGTAACTGCTGCTTGGCCTACTTGAACCTGCGCACGG	24931
QY	177	GGTCAACACATTTTGGCTATCCCGGCGGGCAATTTGGCCATCTATGATGAAGTGA	236
DB	24930	CGTTAGATATATCTTTGGTATCTGCTGGGGCGATCTGCCGATTATGATGATCTGTA	24871
QY	237	CCGCTTTGAAGCGGGGGGAAATTGACCATATTTTGGTGGCCATGACAAAGAGCTTC	296
DB	24870	TAAAGTGAAGCACTGGACGCTTAAACATATTTCTAGTGAACAAGCAAGAGAGAGC	24811
QY	297	CCATGCGGCGGATGGGTATGCGACAGCCAGTAAAGTGGAGTTGTTCCGTACATC	356
DB	24810	TCATCTGCTGATGGTTATGCTGCTGCAACCGGAAAGTAGTGATGCTTTGGTACTTC	24751
QY	357	TGACACAGGGGAGCTAATCTTGGTGAACCGGCAATGGCCATGCCATTTGACTCGGTGCC	416
DB	24750	TGGGCTTGGGCACTAACTTGTGTAAAGGTATGCTACAGCTTACATGATTCATATCC	24691
QY	417	CATGTGGTGAATTAATGAGAGTGGGCGTGGCCATGATGTGAGCATGCTTCCAGGA	476
DB	24630	AATGTGATTTGTTACTTGGGCAAGTACCCCGAAATGATCGTACAGATGCTTCCAGA	24631
QY	477	AATGACATTTTGGCATCACTTACCTTACCGATCGTTAAGCATCTATGTGTAGCTAGTGC	536
DB	24630	AACCGATATCTAAGGTATTAATCTTACCTTACCTTGTAAAGCATCTTATGATGAGTACC	24571
QY	537	GGCGGATATGCTGCGCATTTGTTACTGAGGCTTTCCATCTTGTGACACCGGTGCTCCGG	596
DB	24570	CCMAACATGGGCGCAATTTGGCAGAAAGCATTCACATPACGACACAGGCGGTCCAGG	24511
QY	597	GGCGGTTTGAATGATATTTCCCAAGATGTGGGCTTAAAGAAATGTAGTACATTTCCCT	656
DB	24510	GCGATGTTGATGATGTGCGGAAGATGTGGCTTTGAACAATTTGATTAATGACCTGT	24451
QY	657	CGACCCCGGTGACGTTATCTACCGGGTTATCGCCCGACGGTTAAAGTATCCCGACA	716
DB	24450	AGAACACAGGTTCAATTAATTAACGGGATATCGCCCAAGGTGAAGGAAATTCACGCCA	24391
QY	717	AATTAATCGGCAATGCAATTTTGGACAGAGCCAGAAATCCCTGTGCTACGTAGGAGG	776
DB	24390	AATTAACCCCGGATGATGATGATTCGGGAAAGCCGTGTGCTTTACTATACGTGGTGG	24331
QY	777	AGGGGCGATCGCGCAATGCCCATGCGCAAGTGAAGAAATTTGGGAAAGTTCCAGTT	836
DB	24330	TGGGCGATCGCATTCGCGCTCATGCAAAATTCAGAACTCGCGCGATTAATTAAT	24271
QY	837	GGCGGTAACACCACTGATGGGAAATTTGGGCTTTTGAAGAAACATCCCGCTTGGGT	896
DB	24270	CCCGCTACTACACCTTAAGGATTCGGGGCTTTTATGAAACCAATCTTTATCTTT	24211
QY	897	GGGATGTTGGATGATGCGCACCGGTATGCCAATTTCCGTCAAGCATGTGATTT	956

DB	24210	GGGAGATGTTGGGGATGACGGCACAGCTTAACGTTAGCTGTGACTGATTTGATTT	24151
QY	957	GTTGATTTGCAATGGGGGCCCGTTTCAGCAGACCGGGTAACTGACAACTAGACGAATTTGC	1016
DB	24150	ATTGATTTGGGTGGGTGCTAAGTTTATGATACCGGTGTTACTGCGAAATTAATGATGATTC	24091
QY	1017	TACCCGCGCAAGTAACTTCACTTGAATGACATGCAACCCGGGGAGTGGGAAAAACAGGGC	1076
DB	24090	TTCTGTGCGCAAGTAACTTCACTTGAATGACATGCAACCCGGGGAGTGGGAAAAACAGGCAT	24031
QY	1077	TCCCGATGTCCTCATTTGGGGGATGTAGCCCATGTTTGAACAGCTTTTGACGCGGCG	1136
DB	24030	TCTTGAAGTCCCATTTGTTGGGATGTGGGAAAGTCTTACTGATTTATGATGCTGTTG	23971
QY	1137	CCGGGAATTTGATTTACCCATCCCATACCAACCCAGCATGTTTAAATCGCATTTGA	1196
DB	23970	CAAAACAACAGTGGCCAAATCAACCCCAACCAAAACCAAGAAATGCTTAACTTAATTA	23911
QY	1197	TCAATGGCGGACCGATTTACCCCTCCAGGTGCCCATATGAGGATTAATGCCCCCA	1256
DB	23910	CGTTGGCGGAGATTAACCCCTAGTATGTTCCCGACATCCAGACATTTCCCCACA	23851
QY	1257	GGAGTATGACAGAAATTTGTGCGCAGGCCCCCGATGCTTACTACACACCGATGTGG	1316
DB	23850	AGAAATGATTTGGAAGTGGTATGCAAGCCCCCGCTTCTACACCATGATTTGG	23791
QY	1317	ACAAACCAAAATGTGGGCGGCGGAGTTTGAACAAATGGCCCCCGGATGATTTCCAG	1376
DB	23790	TCAACATCAAAATGTGGGCGGCGGAGTTTGAACAAATGGCCCCCGGATGATTTCCAG	23731
QY	1377	TGCTGCTTGGTACATGAGGCTTTGTTTACTGCTGCGCATGAGGACCAAGTGGAGT	1436
DB	23730	CGAGGCTTGGGAACTATGAGGTTTGTGTTTACTGCTGCGCATGAGGACCAAGTGGAGT	23671
QY	1437	GGGGGACAGC -GGTATTTGATCACTGATGAAGTGGCAGCTTCCAAATGTAATCTTACGA	1495
DB	23670	CCAGATGAAACAGTATCTGATTTGCGGTATGCGATTTCCAAATGTCCTGACAA	23611
QY	1496	ACTGGAAACCTTACCGCCAGTACAGCATTCAGGTTAACTATTAATCTCAATACGGTTG	1555
DB	23610	GCTTGTGATCTGCTGACATATTAAGTAAATGTCAAGACTGTAATCTTAATTAACGGCTG	23551
QY	1556	GAGGGGATGTGCTGACGTGACGACAACTTTTCAAGAAAGCTTATTTGCTTCTAA	1615
DB	23550	GCAAGGATGTGACGCAATGACCAACAGATTCACGCGGAGGCTACTCCTGCTCAAA	23491
QY	1616	CATGTCCAGGCGATGCGACATTAATCTCCTGTGTAAGGCTATGAGCATCAAGGATAT	1675
DB	23490	TATGAAAGTAAAGGATGCGAATGTAAGTATTTGGACAAAGCTTACGCGCATTAAGGATAT	23431
QY	1676	TACTGTGCGAAGCGGAAATTTGGCCCGCGGATGCGGAAATGCTATGCCCAATGG	1735
DB	23430	TGTAGCACTCACAGAGAAATGAAAGATGGGATGCTCAAAATGTTAGCTTACATGG	23371
QY	1736	TCTGTGTGATGAGTGTGTGTCAAAAAGTAAACTGTATACCTATGATTTGCCCC	1795
DB	23370	CCAGTATTTTGAATGTTGCGGTCAAAAAGCGAAATGTGATCTCATGATGATCTCC	23311
QY	1796	CGGATGAGTAATGCCCAATGCTAGTTTACCGAAAGTGGCGGTACNAGACATAGTCC	1855
DB	23310	TGTTAAGATTAATGCTCAATATGATTTGTTTACGAA ---GCAACCAACCAAGCAGTGT	23254
QY	1856	CCGAGTGTGAGTGTACCATTTGCCAAACCCAAATTTCAATCACCATCGTTTCTGTTG	1915
DB	23253	CGAACCAATTTATTTGACCATTTGTGCAAGTCAAAATTTACCAATCACAATTTCTGTTG	23194
QY	1916	TGTTGTGAGGCAACTCTAACCCATTAAGCCAAATTTGAATTT	1958
DB	23193	TGAATGTGGACTAATGTTGTAATTTCTTTGCACTTATTAAGT	23151

RESULT 5
 SPUNHASL

LOCUS SPURASL 1737 bp DNA linear BCT 26-Apr-1993
 DEFINITION *Spirulina platensis* acetylhydroxy acid synthase (ilvY), 5' end.
 ACCESSION M75907
 VERSION M75907.1 GI:152906
 KEYWORDS acetylhydroxy acid synthetase.
 SOURCE *Spirulina platensis*
 ORGANISM Bacteria; Cyanobacteria; Oscillatoriales; *Spirulina*.
 REFERENCE 1 (bases 1 to 1737)
 AUTHORS Milano, A., De Rossi, E., Zanaria, E., Barbierato, L., Ciferri, O. and Ricciardi, G.
 TITLE Molecular characterization of the genes encoding acetylhydroxy acid synthase in the cyanobacterium *Spirulina platensis*
 JOURNAL J. Gen. Microbiol. 138 (Pt 7), 1399-1408 (1992)
 MEDLINE 92381487
 PUBMED 1512571
 COMMENT Original source text: *Spirulina platensis* (strain CI) DNA.
 FEATURES
 source
 Location/Qualifiers
 1..1737
 /organism="Spirulina platensis"
 /mol_type="genomic DNA"
 /strain="CI"
 /db_xref="taxon:1156"
 1..1737
 /gene="ilvY"
 1..>1737
 /gene="ilvY"
 /EC_number="4.1.3.18"
 /codon_start=1
 /product="acetylhydroxy acid synthase (AHAS)"
 /protein_id="AA26595.1"
 /db_xref="GI:152907"
 /transl_table=1
 /translation="MODKQAVKRVTAFAFLIDSLRRHGVQHIFGYPGSNUPIYDEI
 YRAQAGEIKHYLVREHQAHAADGAVSTGKAGVCLATGSLVLAFAHFLM
 SDPLATITGVPRALGTDAFOEIDIFGLPIVKISLYRSESLPRIYVAFHLD
 SGPRGPLIDIPKDVNAOIDYIPEBGSRRVRYPTBENGROINOLQIISLSTK
 PLUYGGGALMAAHAIASELSERFQIPVSTLMGKRPRBNPISLIGVIMGHGT
 AYANFAVMEJDFVAVGVRFDVAGDGFASAXVHIDIPAEVGKSTDVIV
 GDVAVOYIGMLQRTYVHERKLSRNKPNNGIDLQLEPILIVPHEPDGISPDGME
 LSHQCPDAFTYHVGQHWMAQGFVQNGPRRMTSGGLTSGYGLAAVGVKAFHD
 TVTICSDGSEFQMMOMDELGTIAQYIGIKVYIINNGMLGVROMCHMFYNDREATNL
 EDGTEPFAFLADYVGLAEAMVVRKIVQRRLPALSHKGMILDVAVTDEDCYPMVA
 PGHNDSMMLSS"

ORIGIN
 Query Match 34.9%; Score 683.4; DB 1; Length 1737;
 Best Local Similarity 64.1%; Pred. No. 1.2e-188;
 Matches 1097; Conservative 0; Mismatches 601; Indels 13; Gaps 4;

QY 128 AACGGCAACCGGGGCTTATATCTGATGATAGCTGAAACGCCATGGGGTCAAAACCA 187
 DB 26 AAAGGTAACAGAGAGCCCTTGCTGATGATAGCTGCGCCGCAAGGTGACAGATA 85
 QY 188 TTTTGGCTATCCCGCGCGGCAATTTTGCCTATGATGAAGTGTACCGCTTTGAAG 247
 DB 86 TTTTGGCTATCCAGGAGGTTCAAACTTACCATTTGCAAGAAATATAGAGCGGAA 145
 QY 248 CGCGCGGGGAAATTGAGCATATTTTGGTGGCCATGAACAAGAGCTCCCATGGCGCG 307
 DB 146 AAGCGGGGAAATAAACATTATCTGTGCGGCAAGAAAGAGCGCCCAAGCTGCGG 205
 QY 308 ATGGGATGACGACGACAGGTAAAGTGGAGTTTCTTCGGTACCTTGACACAGGG 367
 DB 206 ATGATATGCCCGGAGATACCGGAAGGTGGGGGATGTCTAGCCACTTCTGACCGGAG 265
 QY 368 CGACTAATTTGATGACCGGCAATTGCAATGCCATTGGACTGGTGGCCATGGTGTGA 427
 DB 266 CCACCAATTTAGTACAGGTCTAGCCACTGCTTATCTGATCTGTCGCCGTTTGACA 325
 QY 428 TTACTGAGAGGTGGCGCGTGCATGATGGTGTAGCGATGCTTCCGAGAAATTGACATT 487
 DB 326 TTACCGACAGTCCCGCCGACGCGCTTGGGAACAGATGCGTTTCAGAAATTGATATT 385

QY 488 TTGGCATCACTTACCCGATCGTTAAGCACTCCATGTGTGATAGTGGCGGATATAG 547
 DB 386 TTGGGATTAACCTTACCACTTGTAAACATTCCTATCTAGTCCGGAACCTTGGAAATTAC 445
 QY 548 CTGCAATTTGATGAGGCTTTCCATCTGTGTAGACACGGTGTCTCCGCGGCGTTTGA 607
 DB 446 CGAGAAATTTGTGTGAGGCACTTTCACCTAGCCATGAGTGGCAACCGGACCTGTGTAA 505
 QY 608 TCGATATTTCCCAAGATGTGGCTTGAAGAAATGTAGTATATTCCTCCGACCCCGGT 667
 DB 506 TTGACATTTCCCAAGATGTGGTAAATGCTCAAAATGATTAATTCCTGTAAACCAAGGT 565
 QY 668 AGGTAAATCTACCGGGTATCGCCCAAGTAAAGTATCCCGCAAAATTAATGCG 727
 DB 566 CAGTCAAGCGGTGGGATATGTCTCCACGAGCGGGGAAATCTCGACAAATTAATGAG 625
 QY 728 CATTGCAATTTGTGAGACAGGCGAATAATCCCTTGTCTTACGTAGGGGAGGCGATCG 787
 DB 626 CTTTACAGTTGATTTCCAGACCTCTAAACCCCTCTATATGTGGGGGAGGTCAATTA 685
 QY 788 CCGCAATGCGCATGCGCAGGTCAGAAATTTGGGAAAGGTTCCAGTTGCGGTAAACA 847
 DB 686 TGGCGGGGCTCATCTGAAATTCGCGAATTTGTGGAAACGCTTCAATATCAATTA 745
 QY 848 CCACCTGATGGAAATTTGGGCTTTTGAAGAAACATCCCTTTTCTCTCTCTCTCTCT 898
 DB 746 GCACCTTATGGGAAAGGCTGTTTGTGAAATCAACCCCTCTATCTCTGGAAATCTGTG 805
 QY 899 GATGTTGGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 958
 DB 806 GATGTTGGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 865
 QY 959 TGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1018
 DB 866 TGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 925
 QY 1019 GCGGCGCCAAAGTATTTCAATTTGATGATGATGATGATGATGATGATGATGATGAT 1078
 DB 926 ATAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 985
 QY 1079 CCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1138
 DB 986 CTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1044
 QY 1139 GGGAAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1198
 DB 1045 TATCATTTGGAACGGAATTTAAGCCGGAACCAACCCAGGATGG--CACTGACCTTATC 1102
 QY 1159 ATTGCGGACCGATTAACCCCTCCAGGTCGCCCTATGAGATTAATTTCCGCCAGG 1258
 DB 1103 AATTGCGGAACCAATATCCCTTGACGTTCTTCAACCGGAAGATGATGATGATGATGAT 1162
 QY 1259 AGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1318
 DB 1163 ATGCGATTTGGAATTTAGCATCAATGCTCTGATGATGATGATGATGATGATGATGAT 1222
 QY 1319 AAACCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1378
 DB 1223 AGCATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1282
 QY 1379 CTGCGTTGGTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1438
 DB 1283 GCGGTTTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1342
 QY 1439 GGGAGCA-GCGGTATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1497
 DB 1343 CTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1402
 QY 1498 TGGGAACCTTACCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1557
 DB 1403 TGGGAACCTTACCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1462

QY	1558	AGGGATGTCGTCGTCAGTGGCAACAACCTTCTACGAAGACGTTATTCGCTTCTACGA	1617		
Db	1463	TGGGAATGTCGTCGTCAGTGGCAACAACCTTCTATATATGACCGCTTACAGGCTACTAAC	1522		
QY	1618	TGTCGCCAGGCGATGCGCACATTAATCTCTCTGTGAAGCCTATGATCAGAGGCTATTA	1677		
Db	1523	TGGAAGATGGAACCCCAATTTGCGCGTTAGCTGATGTTTATGACTTGAGGCTATGA	1582		
QY	1678	CTGTGCGCAAGCGGGAATTTGGCCCCGCGATGCGCGAATGCTAGCCCAATGATC	1737		
Db	1583	ACGTCGCGCAACGGAAGATTTACCAAGGCGACTGCTTAAGCTCTATCTCACAAGGATC	1642		
QY	1738	CTGTGTGATGATGATGTCGTCGTCAGTGGCAACAACCTTACCTATGATGCCCCCG	1797		
Db	1643	CGATGATCTCTAGATGTCGCGTTAACCGTGATGAGAGACTGTACCAGATGAGAGCCCCG	1702		
QY	1798	GCATGAGTAAATGCCCAATGCTAGGTTTACC	1828		
Db	1703	GTCACGATTAACAGCGACATGATGCGTTTATC	1733		
RESULT 6	AP006579	297800 bp	DNA	linear	BCT 18-OCT-2003
LOCUS	AP006579	Gloeobacter violaceus PCC 7421 DNA, complete genome, section 12/16.			
ACCESSION	AP006579	BA000045			
VERSION	AP006579.1	GI:35213671			
KEYWORDS					
SOURCE					
ORGANISM					
REFERENCE					
AUTHORS	1	Gloeobacter violaceus PCC 7421			
		Bacteria; Cyanobacteria; Chroococcales; Gloeobacter.			
TITLE	2	Nakamura, Y., Kaneko, T., Sato, S., Mimuro, M., Miyashita, H., Tsuchiya, T., Saeemoto, S., Matsumoto, M., Kawashima, K., Kishida, Y., Kiyokawa, C., Kohara, M., Matsumoto, M., Matsuno, A., Nakazaki, N., Shimp, S., Takeuchi, C., Yamada, M. and Tabata, S.			
JOURNAL		Complete genome structure of Gloeobacter violaceus PCC 7421, a cyanobacterium that lacks thylakoids (supplement)			
REFERENCE		DNA Res. 10, 181-201 (2003)			
AUTHORS	3	Nakamura, Y., Kaneko, T., Sato, S., Mimuro, M., Miyashita, H., Tsuchiya, T., Saeemoto, S., Matsumoto, M., Kawashima, K., Kishida, Y., Kiyokawa, C., Kohara, M., Matsumoto, M., Matsuno, A., Nakazaki, N., Shimp, S., Takeuchi, C., Yamada, M. and Tabata, S.			
TITLE	4	Complete genome structure of Gloeobacter violaceus PCC 7421, a cyanobacterium that lacks thylakoids			
JOURNAL		DNA Res. 10, 137-145 (2003)			
REFERENCE					
AUTHORS	5	Kaneko, T.			
TITLE	6	Direct Submission			
JOURNAL		Submitted (15-AUG-2003) Takakazu Kaneko, Kazusa DNA Research Institute, The First Laboratory for Plant Gene Research, 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba 292-0818, Japan			
FEATURES		(E-mail: kaneko@kazusa.or.jp, URL: http://www.kazusa.or.jp/cyano/, Tel: 81-438-52-3935 (ex.2338), Fax: 81-438-52-3934)			
SOURCE		Location/Qualifiers			
gene	1..297800	/organism="Gloeobacter violaceus PCC 7421"			
CDS		/mol_type="genomic DNA"			
		/strain="PCC 7421"			
		/db_xref="taxon:251221"			
		complement(7..948)			
		/gene="argP"			
		complement(7..948)			
		/gene="argP"			
		/note="ORF ID:g113101"			
		/codon_start=1			
		/transl_table=1			
		/product="ornithine carbamoyltransferase"			
		/protein_id="BAC91042.1"			
		/db_xref="GI:35213672"			
		/translation="MSASLGAATFRPDLISLDDLEAQLHALITLALQLRGRVANLHGKVLGVFLKASTRTVSVFVAMVQLGGQVILDSPTNTQVGRGEPVTRAVLGRV			
gene		DGLAIRTPAQTELEEVAYAGIPVINALTDHEHPQVAVADLLITRENFRLAGLKLAV			
		VADGNNVAHSLILGCAKVGMSIAVATPEGFTDPVAVASASEIAERTAEVQILNDPF			
		EAARGAHILITVDTVMISMGAEQTHLQIFRQYQYNALINCAAEAIYHCLPAHRG			
		BEITDEVMEGPRSRITWDEAENRILHAOKVLAALMGSR"			
CDS		complement(945..2282)			
gene		/gene="g113102"			
		complement(945..2282)			
		/gene="g113102"			
		/note="ORF ID:g113102"			
		/codon_start=1			
		/transl_table=1			
		/product="carboxyl-terminal protease"			
		/protein_id="BAC91043.1"			
		/db_xref="GI:35213673"			
		/translation="MKTPKFLIPVLGACATVILSLISAFNMPGMANFKAPKEIIDEVW			
		QVNNREYVDPTFNQNMOWETRRKYLAKYKSEDEAYKSTREMLKGLGDPYTFMDPKQ			
		YSKMKVETSGDYQGVQIQLGDEKTHELVVSPIDTAFATGVPRKVLISIDGRST			
		KEMDIDQAVNFRGOAGTSVALKFRREGKFLTLIVRRIELKPKVSLRTEGDRVG			
		YIRLSQFNAYAAKDMGNVAKLTREQADGFTLILSNQGLLYASADIAARLISGTL			
		VSTVDRDGRBESITYSRPMATSKPLVYLVDGASASBELISGALDDHRAVYVGRTY			
		GHGVQVVASLSDSGVAVTIAHYOTPSGRDINKRGIDPDITRELTKEAKVRLKADQV			
		ATLADPQYRAVDVILGQQLAGMPVEQVVAAGKPPPAERQPVAAENQPVAVP"			
		complement(3391..5712)			
		/gene="g113103"			
		complement(3391..5712)			
		/gene="g113103"			
		/note="ORF ID:g113103			
		hypothetical protein"			
		/codon_start=1			
		/transl_table=1			
		/protein_id="BAC91044.1"			
		/db_xref="GI:35213674"			
		/translation="MEAPLDNGCCDLLAIYSRTITVKEPDSFGGCGOLFVNLICKS			
		HPNPKTPKRAKSSISPGGYTPSREKGNLRHTRIKULISMSVRIKRTKGTALVILA			
		GFQLVGLMSAQIIFTRDSLAKPSQAADLAAPVGPVAPQVSPFPDPAADPDGATT			
		LDLETVEGLAARRSTPVTITREDIQRKAAVNLADALRGVGFALINDGRADPANG			
		SVYKQVTVQMSALLNGRQIKTGLSNMGATVNSIPEGIRIELLGFAGCTTLYGS			
		DVIGGVINIIITEEGPAVPTPFGSAEYSGSIVRTSGTSERLKFNFGVQSYCATN			
		NPPVAPVSTNRGDRKLNADMYTASVQSAITPDRNNLTFDLFKLSRGLIIGN			
		SKDRLRDSGLNVLGRTLLIGRNDTSVITTVGVDDEPSTYGPVQDTYFKAVOLASL			
		FYQARDHWRTSPNTLRMGVDTNHTLISQNTTSPDPIFRGCGTTDTATNTNGI			
		FLDITQVIEGQLDGLRQITDSKFSGLNSAGLRKQIIEVALRASVAVQKNGP			
		PDPLPLDAEGLFNPDLKPERGSVNAGPDLRPAADLRGVVTFGLIDRIQNGP			
		LSADPLTQRTNIGVNTGVNAGIKMQIDPQMSYSYTKRIKIEGRSGILRFL			
		PPSVCQVGVNQGLOVNVFVNVNSGAARTALTSYTFPATIDLRALPIFNALAL			
		TYVLANLDVQYEVNTFVSPGLTVRVGLQTSF"			
		5711..6454			
		/gene="g113104"			
		5711..6454			
		/gene="g113104"			
		/note="ORF ID:g113104			
		probable ABC transporter ATP-binding protein"			
		/codon_start=1			
		/transl_table=1			
		/protein_id="BAC91045.1"			
		/db_xref="GI:35213675"			
		/translation="MLEAVDLTKNYGTAVPISLGINLRILEAGEVECLLPNGAKXTT			
		ISLFLGFIAPSAGTARIKSLDVORHPLRTPKRVAYIPQWMLYRMLSGLENELEYSAL			
		AGQDRTYRELLAPPEQVGLPAESAGRGRTGSKPRQKVGIAIITAKAEALIDDEP			
		TSGLDPKASNESELLGRLSRGVAVLMAATHDLPAKXAGRVGVIMNAGRLVTTARTE			
		DLVYQRLRTLYESNVGASAGCAVEVR"			
		6451..7866			
		/gene="g113105"			
		6451..7866			
		/gene="g113105"			
		/note="ORF ID:g113105			
		hypothetical protein"			
		/codon_start=1			
		/transl_table=1			
		/protein_id="BAC91046.1"			
		/db_xref="GI:35213676"			
		/translation="MIMRIARKEVYLENTRDGRFRAGATVVLGLLVVSLMGKRYSDV			
		SAQHAAGQARTQWLQDEPKNHSAAHYGVYAFKPLPLSFVDGQVDVYGVAVVLE			

AKKONDERPAPADATAVORFELTAAVLOLLPLVILLGFGAPAGEREGTLEOL
LSIGVPAATLALGALMVALIGLILVPAITLIGVALALAGGGIILAPLIALISA
AYIGYLLFVGYCLAVASAMAKSARIALVAILGFWIYOGVALPRLVILARGLYPSPA
LTTETVATDMAGGIDGSHNSQDARPALEKRVITQYASREVDLPVNFAGIALQESSE
YGRVYDRRGELMRPEEONAVHOGFVAPLILARSMSJSDGDFQOHPFATA
EOYRGRGIKTMEATMENKDDGRATADRVSDFOYTPPLGWLVAQOOSJLLGLG
FWIYAAAAVFAARLRYTG
7866..9311
/gene="g1r3106"
/note="ORF ID:g1r3106"
unknown protein
/codon_start=1
/transl_table=1
/protein_id="BAC91047.1"
/db_xref="GI:35213677"
/translation="MTASIIKHEWRNLMAERVIMAVGGLPALLIYVGLINGTGWQFO
RQTLAQOQSEITRLDPLQSIADISGKVRFGDAPADROPRIVGNKGRVIALTL
RPLAPLAVQSDLYPYFVSTLANKQTFIÖNNELNPTNLARGDLAFVILPL
FIALGYNLGSRDEBGLTAMLISOPVGLMVLGVLVAGGIVLIALVLSIAGVL
GGADPAEPGRILGLMVLVILGYMFWLAVANAAGRTPTNAVLVGLMLV
IVPAVNVAVSTSPVSRVYLQAVREATNANAGSRLIAYEDHDELAPKDSRV
DGSFALRSTAVRETVDKATAPLERYDROLAQOALVRYVYLSALVIAOHEVDILA
GTGIGRHRFOVDRHRTWQGYFVPRITRQCTVTAELGRLPKRFREBELTGVAG
RVLPGLVGLAPLVLVGLTLPGLRRPVAA"
complement (9377..10342)
/gene="g113107"
complement (9377..10342)
/note="ORF ID:g113107"
hypothetical protein
/codon_start=1
/transl_table=1
/protein_id="BAC91048.1"
/db_xref="GI:35213678"
/translation="MRRIVVALLMMGLATPAPAPAKIGYSAMPQMPMKVABAGL
PKKNGVQWQWVEGVADSTALVAGLNDNCOTLDTTAPAAAGAKLVLTNNSA
GNQOILYKKEIASIKDLKKKGLGELCTVDHYLLIGQKEGLTODTDPETLNSAA
AAAPAGRLDGVGFAPFTTALSTGKAVLFSKKEPFGAISDLVLRBEVKAAPGE
IEKIVKTFPDLTYIKNPESLQIMAKQGSTVAEKYSISGTRLFNLEONLVTFAP
ETQDPINLRYOSKIADFLVKSGLAKTKPTAVLPEDKFRALKP"
10476..10730
/gene="gsr3108"
10476..10730
/note="ORF ID:gsr3108"
unknown protein
/codon_start=1
/transl_table=1
/protein_id="BAC91049.1"
/db_xref="GI:35213679"
/translation="MFIHADLTRHFEETPLPSLPZHRAARTLHAQKRLDALREBS
TGMFTAGQETARAECAKELVDAETEIRAKIVLKAQPO"
complement (10782..11156)
/gene="g113109"
complement (10782..11156)
/note="ORF ID:g113109"
hypothetical protein
/codon_start=1
/transl_table=1
/protein_id="BAC91050.1"
/db_xref="GI:35213680"
/translation="MAKNKRSOTPROIPGLPVPGNRVRLDPOVLDAGRVAVI
DCSASERRIMTLAGVAPDARFEINNRKGVVYCVSMWLSRSIVAAVRLAKGYT
AYDLKGLAAWKTAGYPTIRAL"
11432..11905
/gene="g1r3110"
11432..11905
/note="ORF ID:g1r3110"
unknown protein

Query Match 34.3%; Score 672; DB 1; Length 297800;
Best Local Similarity 62.4%; Pred. No. 5,88-185;
Matches 1127; Conservative 0; Mismatches 656; Indels 22; Gaps 4;
/codon_start=1
/transl_table=1
/protein_id="BAC91051.1"
/db_xref="GI:35213681"
QY 136 ACCGGGGCTTATTCCTGATGATAGCTGAAAGCCATGGGGGCAACAACATTTTGGC 195
DB 181531 ACCGAGGCTTCGCGGTGCTGATAGTCTGCTGCAAGGAGCGTCCGACATCTTGCT 181590
QY 196 TATCCCGGAGGGCAATTTTGCCTATGATGAACTGACCGCTTTGAAGCGCGGG 255
DB 181591 TACCCGGTGGGCAATTTTCCCGGTCTACAGACAGATTAAAGAGCCGAGCGGACTAC 181650
QY 256 GAATTTGAGCATTTTGTGTGCTGCGCCATGAAACAAGACCTTCCCATGGGGGAGATAT 315
DB 181651 CGTTTCAGCACTTTCGTGTCGCGCAGACAGAGAGGGGGGCGACGCGCGGACTAC 181710
QY 316 GCCAGAGCCAGAGTAAAGTGGAGTTGTTTCGTTACATCTGACACAGGGGCGACTAAC 375
DB 181711 GCCCGGCGAGCCGGCGGGGTGGGCTGTGTGCGCACTCCGGTCCCGGGGCGAGCAAT 181770
QY 376 TTGGTGACCGGCATTTGCCAATGCCCATTTGACCTCGGTGCCATGCTGTGATTACTGA 435
DB 181771 CTGGTGACGGGCATCGCCACGCGCCACATGATTCGGTCCGATGAGGCGATTAACGGT 181830
QY 436 GAGGTGGCGCGCATGATTTGTAGCGATCTTTCCAGAAATTTGACATTTTGGCATC 495
DB 181831 CAGGTCCGCGCGCGGTGATGCTGCTCCAGCCCTTTCAAGAGCCGATATTTCCGGCATC 181890
QY 496 ACCTTACCGATGCTTAAAGCACTCTGTATGCTAGCTAGTGTGCGCGGATATGCTCGCAT 555
DB 181891 ACCCGCCCATCGCAAGCACTCTTACGTATCCGCAACTGCGGGGCACTCGCGCATC 181950
QY 556 GTTACTAGGCTTTCATCTTGTAGCAACCGGTGTCGCCGGGCGGCTTTGATGATAT 615
DB 181951 TTACGCAAGCTTTCATATGCAAGCAACGCGCGCGCGCGCGGCTTCATGAGCTT 182010
QY 616 CCCAAGATGTGGGCTTAAAGAAATGTAGTACATTCCTCGACCCCGGTGACGTTAAT 675
DB 182011 CCCAAGAGCTGGGCAAGAAAGCTTCGAGTACGCGCGGTGCAAGACGCGCC---AAA 182067
QY 676 CTACCGGCTTATCGCCCAACGCTTAAAGTAAATCCCGCAAAATTAATGGGCACTTGCA 735
DB 182068 ATTCCCGGTACCGCCCAACCAACAGGAGCATCCCGCAAGTGGCGGCGGCGAAG 182127
QY 736 TTGTTGAGCAGGCGCAAAATCCCTGCTTACGTAGAGGAGGCGGATCGCCCAAT 795
DB 182128 CTGATCCCTGAGAGGAGACCGGCGCTGTGTTATGTGGGTGGCGGCGCATCTCGCGGG 182187
QY 796 GCCATGCGCAGGTGCAAGAAATTTGCGAAAGTTTCCAGTTGCCGTTAAACAACCTCTG 855
DB 182188 GCGCAGCGCGCAACTGATGCGGCTGCGAGTGGGTTTGAATCCGACACACACACCTTG 182247
QY 856 ATGGGAATTTGGGCTTTTGAAGAAACCATCCCTTCCGTTGGGTATGTTGGATGATCAT 915
DB 182248 ATGGGCAAGGGGCGCTTCCCGAAGATCATCCGTGCGCGTGGGCAATCGCATGAC 182307
QY 916 GGCCACCGCTATGCCCACTTTCGCTGACGAGATGATTTGTTGATTCAGTGGGCGCC 975
DB 182308 GGCACCGGCTATGCCCAATTTGCGCTGACGAGTGGCATTTGCTCATTCGCGTGGGCGG 182367
QY 976 CGTTTCGACGACCGGGTAACTGGCAAACTGAGCAATTTGTAGCGCGCGCAAGTAAAT 1035
DB 182368 CGCTTCGACGACCGGGTAACTGGCAAACTGATTAATTTGCGCGCGCGCGCGCTCATC 182427
QY 1036 CACATTTGACATGACCCGCGGAGGTGGGAAAAAACAAGGCGCTCCGATGTGCTCCATTGTG 1095
DB 182428 CACATGATATTCGATTCGGGAGAGGTGGGCAAGAACCGGGGCGGAGGTGCCATGCTG 182487

QY 1096 GGGAGTGTACGCCATGTTTGTAGACAGCTTTGACGCGGCGCCGGGAATGATTAACCC 1155
 Db 182488 GCGGCGCTCAAACTGTGTGAGGATCTGATTAATAAGCTCATGAAATTT----- 182537
 QY 1156 ACCATCCCATACGACCGGATGTTTAATGCGATTTGATTCGCGGACCGGATTAC 1215
 Db 182538 -----TCAGCCGCGGACCGGCGGTGGTGGAGCCCATGAGCGCTGGAAGCGGATAT 182592
 QY 1216 CCGCTCCAGGTGCGCCCATATGAGATTAATGCGATTTGATTCGCGGACCGGATGATCAGCAAT 1275
 Db 182593 CCGCTGAGGTGCGGACGATGAGCAGCGGCTGTATCCACAGGAAGTATGATGCGCTTT 182652
 QY 1276 GGTCCCGGAGCGCCCGGATGCTTACTACACACCGATGTGGGACACACCAATGTGGCG 1335
 Db 182653 CGGAGCTGGCTACCGAGCGCTACTACACCGAGCGTGGCGCAGACCAAGATGTGGCG 182712
 QY 1336 GCCCAGTTTTTGACCAATGCGCCCGCGATGATTTTCAGTGTGCTTGGATACGATG 1395
 Db 182713 GCGCAGTTTCTCAAAAAGGCGCGCGGCTGTGATTTCTCAGCGGCGCTCGCACGATG 182772
 QY 1396 GCGTTTGGTTTACCTGCGCCATGAGGACCAAGTGGAGTGGGGACG-AGCGTCAAT 1454
 Db 182773 GCGTACGCGCTTCCGCGCGCATCGCGGTGAGTGGACCTGACGAAAGCTCGTG 182832
 QY 1455 TGCACTAGTGGAGATGCCAGCTTCGAAATGATCTTCAGGAAGTGGGAACCTAGCCAG 1514
 Db 182833 TGCACTCCGCGGACGCGGCTTTGATGAACTTACAGAAATGGCAGCGCGTGCAG 182892
 QY 1515 TACGACATCCAGTTTAACTATTAATCTCAATTAACGATGGGAGATGTGTGATG 1574
 Db 182893 TACGACATCCAGTTTCAAGGCGCGCATCTGTCAACAGCGCTGCGAGGATGTGCGGAG 182952
 QY 1575 TGGCAGCAAACTTTTATAGGAAGAGTATTTCTGTTTAACTATGCTCCAGGATGCA 1634
 Db 182953 TGGCAGCAAACTTTTATAGGACCGCGCTTCCCATCTCCACATGAAGCCCTCGATGCC 183012
 QY 1635 GACATTAATCTCTCTGTGAAGCTTATGACATCAAGGATTAATCTGTCCGACAGGAA 1694
 Db 183013 GACTTCGTCAAACTGCTGAAGCGTTGGGGGAGGAGGATGCTATGATGATCAGCCGAC 183072
 QY 1695 GATTGGCGCCCGGCGATGCGGCAATGCTAGCCCAATGCTCTGTGTGATGATG 1754
 Db 183073 CAGCTCAAGAGAGGATGAGGCGCTGATGAGCGCGGCGCGCTGATTTGCGACTTT 183132
 QY 1755 GTGTGCAAAAAAGATGAATACTGTATACCTTATGATGCGCCCGGATGATTAATGCCAA 1814
 Db 183133 CATGTGCTGCGGAGCGCACTGCTACCGGATGCTGCGCGCGGCGGCAAGCCACACAG 183192
 QY 1815 ATGTAGGTTTACCGGAAGTCCGCTACNGACAATGCTCCCGGATGATGATGCAAC 1874
 Db 183193 ATGTGCGGCGTCC---CAGGCGCTTGCCTGACGCGCTGCAACTCATCTATGCCCC 183249
 QY 1875 CATTCGCAAAACCAAAATTTATATACCATGCTTCTGTTGTTGTGGAGCCAACTC 1934
 Db 183250 AACTCGGACCAAAATAACACGCGCATCTGCCCCGAGTGGGACCAAGCTT 183309
 QY 1935 TAAAC 1939
 Db 183310 TAGCC 183314

RESULT 7
 BX569693/c BX569693 349213 bp DNA circular BCT 23-SEP-2003
 LOCUS Synecchococcus sp. WH8102 complete genome; segment 5/7.
 DEFINITION BX569693 BX548020
 ACCESSION BX569693 1 GI:33638930
 VERSION
 KEYWORDS complete genomes: WH 8102
 SOURCE Synecchococcus sp. WH 8102
 ORGANISM Synecchococcus sp. WH 8102
 Bacteria; Cyanobacteria; Chroococcales; Synecchococcus.
 REFERENCE 1 (bases 1 to 349213)
 AUTHORS Patenik,B., Brahamsha,B., Larimer,F., Land,M., Hauser,L., Chain,P.,

FEATURES
 source
 Lamerdin,J., Regala,W., Allen,E.A., McCarren,J., Paulsen,I.,
 Dufresne,A., Partensky,F., Webb,E. and Waterbury,J.
 The genome of a motile marine Synecchococcus
 Nature 424 (6952), 1037-1042 (2003)
 JOURNAL
 MEDLINE
 PUBMED
 12917641
 2 (bases 1 to 349213)
 AUTHOR Larimer,F. and Patenik,B.
 CONSRTM Synecchococcus genome consortium
 TITLE Direct Submission
 JOURNAL Submitted (01-JUL-2003) Larimer, F., Submitted on behalf of the
 Synecchococcus genome consortium, the DOE Joint Genome Institute,
 Production Genomics Facility, 2800 Mitchell Drive, Walnut Creek, CA
 94598, USA, and the Genome Analysis Group, Oak Ridge National
 Laboratory, 1060 Commerce Park Drive, Oak Ridge, TN 37831, USA;
 larimerf@ornl.gov

location/Qualifiers
 1..349213
 /organism="Synecchococcus sp. WH 8102"
 /mol_type="genomic DNA"
 /strain="WH8102"
 /db_xref="taxon:84588"
 /complement(52..1029)
 /locus_tag="SYNW1422"
 /complement(52..1029)
 /locus_tag="SYNW1422"
 /note="Similar to selenophosphate synthetase"
 /codon_start=1
 /evidence=not experimental
 /transl_table=11
 /product="conserved hypothetical protein"
 /protein_id="CAE07937.1"
 /db_xref="GI:33638931"
 /translation="MSFTLMDGLRLHSLGKLAKRDIOPAGNFCQHPDQLGAGMLG
 DPAAL,LRQNGQLLACBQOPGLVEEDWPAWGSGLVNLSDIAMGRPIALVNSV
 WSGADALQPLLEGMFACSEFGVWGGSHSQOQSYTLGVAVGAGPYLSARA
 APDE,LMVNNKGAIFYRHPMDATHTPAPRLREHSL,LPALAEGLVRAKIDIM
 GGLTSTWMEFAEGLESLDIDAYDRPQGVQEWL,SGPSPGYLADLPBRTSLV
 RMLQSPDLICRIGRFSTGBCRVLIQISGASHCLMEVNASGLTRGCGG"
 /complement(1026..1697)
 /locus_tag="SYNW1423"
 /complement(1026..1697)
 /locus_tag="SYNW1423"
 /codon_start=1
 /evidence=not experimental
 /transl_table=11
 /product="conserved hypothetical protein"
 /protein_id="CAE07938.1"
 /db_xref="GI:33638932"
 /translation="WVFCUDPSPSRIGRSISSAPSLFTPSVRAIGIDDPRLSP
 SSRRSFHLIRPDSPLIAGWLSIRSI,FCGEHVFPSQSDRLDRACIALIHHDC
 PBDKGPSSQVGVYRIYETPRMLYGRGLGVSHDFRHNQIGGLIKRATITANG
 GCDRFAMVQIQVRFRLHVASIIEELRIGRLHMDADNYTAPSKRRPSCSL
 PSLAA"
 /complement(1700..2791)
 /locus_tag="SYNW1424"
 /complement(1700..2791)
 /locus_tag="SYNW1424"
 /note="Biotin synthetase-like domain containing protein"
 /codon_start=1
 /evidence=not experimental
 /transl_table=11
 /product="conserved hypothetical protein"
 /protein_id="CAE07939.1"
 /db_xref="GI:33638933"
 /translation="MSERGRITVLEOVKVRAPLEPGRGREGAGSPDRALNDGT
 TWWPYRNVSSSAVSLARBDGLITRPODAPLVYTTTDEPFIQISTREG,PIK
 STLLHSRVLAITLLQCTIRFDRSQSCQFCALIQSDGATVIRKPEVOVAEAE
 AVRLDGVKQLVMTGTPNSDRGARLMEATAAVRVLD,PIQGGCEPDDPIIWERM
 KAAGIDSLGMLHEVELEVYRRIILPGKSELSELYAAFADAVAFGGEVSTYLLAG
 LQDSRALDSCRRILIEGVYPFVYPIISGTPLEHHAPDPAFTAMVDVYKGVALLH
 QDIREAMAGACGACGASALFEGAA"
 /complement(2784..3800)
 gene
 CDS

QY 496 ACCTTACCGATGTTAAGCACTCTATGTGTAGTACGCGCGGATATGCGCAT 555
DB 286884 ACCCTCCCATCGTGAAGCATCTCTGGTGTGCGGACCTCTGTGACCTGGTTCGATC 286825
QY 556 GTTACTGAGGCTTTTCATCTTCTAGACACCGGTCTGTCCCGGCGGTTTGTATCAT 615
DB 286824 GTGGCCAGGCGTTTGTATGCGCCGACGCGGTGCGCCCGCGCTGTGTATGACATC 286765
QY 616 CCCAAGATGAGGCTTGAAGATATGATACATTTCCCTCGACCCCGGTAGCTTAT 675
DB 286764 CCCAAGATGAGGCGGAGAACATCTACATGTCGCGGTGGAACCGGTTCCTCATC 286705
QY 676 CTACGCGGTTATGCCCCACGCTTAAAGTAAATCCCGACAAATTAATGCGGATTGCA 735
DB 286704 CCGTGGCGGTTTCATTCAGCCCAACCTCGCTGATGCTGCGGTGACGCGGCACTGGAT 286645
QY 736 TTGTTGAGCAGCAGCCGAAATCCCTTGTCTCTAAGTGGGGGAGGGGCGATGCGCGCAT 795
DB 286644 CTGATCGAGCAGGCGCCCAAGACCGCTGTCTGTACGTGGGTGCGGGCGATCTCTGCTGT 286585
QY 796 GCCCATGCCAGGTGAGAAATTTGGGAAAGGTTCCAGTTGCCGTTAAACCAACCCCTG 855
DB 286584 GCCCAGCAGAGCTGCGGATGCTGACCGGATTCAGCTCCCGTACACCAACCTTG 286525
QY 856 ATGGGAAATGGGCTTTTGAAGAAACATCCCTTGGGTGGTATGTTGGTATGAT 915
DB 286524 ATGGGAAAGGTGCTTTGATGAAACGATCTCTCTGCGTGGGATCTCGTATGAT 286465
QY 916 GGCCACCGCTATGCCAATCTTGGCGGTACGATGTAATTTGTTGTTGACAGTGGGCGCC 975
DB 286464 GGCCACCGCTATGCCAATCTTGGCGGTACGATGTAATTTGTTGTTGACAGTGGGCGCC 286405
QY 976 CGTTTCAGACGACCGGTTAATCTGCAACTAGCGAATTTGCTAGCCGCGCCAAAGTAT 1035
DB 286404 CGTTTCAGACGACCGGTTAATCTGCAACTAGCGAATTTGCTAGCCGCGCCAAAGTAT 286345
QY 1036 CACATTTGACATGACCCCGGAGGTGGGAAAAAAGAGGCTCCCGATGTGCCATTTG 1095
DB 286344 CACATTTGACATGACCCCGGAGGTGGGAAAAAAGAGGCTCCCGATGTGCCATTTG 286285
QY 1096 GGGGATGTAAGCCGATGTTTGAACAGCTTTTGACGCGGCGCCGGAATTTGATTAACCC 1155
DB 286284 GGGGATGTAAGCCGATGTTTGAACAGCTTTTGACGCGGCGCCGGAATTTGATTAACCC 286234
QY 1156 ACCCATGCCAGTACCAACGAGCATGTTTAAATGCAATGATTCATGCGGACCGATTC 1215
DB 286233 ACCCATGCCAGTACCAACGAGCATGTTTAAATGCAATGATTCATGCGGACCGATTC 286174
QY 1216 CCGCTCCAGGTGCCCCCATATGAGATCTATTTGCCCGCCCAAGAGTATGACAGAAAT 1275
DB 286173 CCGCTCCAGGTGCCCCCATATGAGATCTATTTGCCCGCCCAAGAGTATGACAGAAAT 286114
QY 1276 GGTGCCGAGGCGCCCGGATGCTACTACACACCGATGTTGGACACACCAATTTGGCG 1335
DB 286113 GGTGCCGAGGCGCCCGGATGCTACTACACACCGATGTTGGACACACCAATTTGGCG 286054
QY 1336 GCCCAGTTTGAACATGAGCCCGCGCGATGATTTCCAGTGTGCTTGGTACGAT 1395
DB 286053 GCCCAGTTTGAACATGAGCCCGCGCGATGATTTCCAGTGTGCTTGGTACGAT 285994
QY 1396 GCGTTTGGTTTACCTGCGCCCATGAGGACCAAGTGGGAGTGGGGGA CGAGCGTCAAT 1454
DB 285993 GCGTTTGGTTTACCTGCGCCCATGAGGACCAAGTGGGAGTGGGGGA CGAGCGTCAAT 285934
QY 1455 TGCATCAGTGAAGTACAGCTTCCAAATGATCTTACAGAACTGGGAAACCTTACGCCAG 1514
DB 285933 TGCATCAGTGAAGTACAGCTTCCAAATGATCTTACAGAACTGGGAAACCTTACGCCAG 285874
QY 1515 TAGCAGATCCAGGTTAAATCTATTTCTCATATAAGGTTGGCAGGGGATGTTGCTGAG 1574
DB 285873 TAGCAGATCCAGGTTAAATCTATTTCTCATATAAGGTTGGCAGGGGATGTTGCTGAG 285814
QY 1575 TGGCAACAACTTTCTACGAAGAGTATTTCTGTTCTTAACATGTCCAGGCGATGCA 1634

DB 285813 TGGCAGGAAAGCTTCTACGACGAGCGCTATTTCCGCCCTCGACATGCTCATGCTCCT 285754
QY 1635 GACATTAATCTCTCTCTGTGAAGCTTATGACATCAAGGATTTACTGTGCCAAAGCGGAA 1694
DB 285753 GATTTCATCGCCCTGAGCCCGCTTCTCTGCGGTGATGCGCTCAAGATCACCGACCTGAG 285694
QY 1695 GATTTCATCGCCCGGATGCGGAAATGCTTACCTTATGATGTTGCGGATGATGATG 1754
DB 285693 TTGCTGACACCGGACCTCTGTCAGCCCTGATCACCACCCGACGATGATGATG 285634
QY 1755 GTGTCAAAAAAGATGAAAATCTTACCTTATGATGTTGCGGCGATGATGATG 1814
DB 285633 CACGTGCGGCGGAGAGAACTCTACCCGATGATGTCCTCCGCGGCAAGACAAATGACAG 285574
QY 1815 ATGCTAGTTTACCGGAAGTCCCGGTACGACAAATGCTCC 1856
DB 285573 ATGCTAGTTTACCGGAAGTCCCGGTACGACAAATGCTCC 285532
RESULT 8
PSP224709
LOCUS
DEFINITION Porphyridium sp. chloroplast gene encoding acetylhydroxy acid synthase.
ACCESSION AJ224709.1 GI:3157427
VERSION AJ224709.1 GI:3157427
KEYWORDS acetylhydroxy acid synthase, ahas gene.
SOURCE chloroplast Porphyridium sp.
ORGANISM Eukaryote; Rhodophyta; Bangiophyceae; Porphyridiales; Porphyridiaceae; Porphyridium.
REFERENCE 1
AUTHORS Lapidot, M., Ravesh, D., Sivan, A., Arad, S. and Shapira, M.
TITLE The chloroplast-encoded Ahas gene of Porphyridium sp. is distinct from the nuclear-encoded gene of green algae and plants
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 2338)
TITLE Direct Submission
JOURNAL Submitted (23-FEB-1998) Ravesh D., Life Sciences, Ben Gurion University of the Negev, P.O.Box 653 Beersheba, 84105, ISRAEL
FEATURES
source
1.2338
/organism="Porphyridium sp."
/organism="Porphyridium sp."
/mol_type="genomic DNA"
/db_xref="taxon:77580"
/note="rhodophytes"
544..2307
/gene="ahas"
544..2307
/gene="ahas"
/codon_start=1
/transl_table=11
/product="acetylhydroxy acid synthase"
/protein_id="CAA12081.1"
/db_xref="GI:3157428"
/db_xref="GCA:063181"
/db_xref="SPTREMBL:063181"
/translation="MTHIRKSNVYQETGAFALDLSVRHKVHIFGYPGATILPIYDE
LYKMEQGIKHLVHEQGAHAADGVRATGVEGCPATSGPATNLVTGIATAM
DSIPYITIGOVRSPIGTDAFOEVDIFGTTPIVGSYVIDPRIPRIIVAEPSIA
KGRPEPVLIDVKNVGLTFEYOVVNPPEARIPGRDVAAPSSQIHSIOLIEAN
OPLVYGGATSGAHDLYKYNOKYPIITTTIMKGTIDPQNPLALGMLGMEHAY
ANFAVECHULLTLGRFDRVTGKIDEPACAKYIHDVIDAEVKNIPQVATVGD
ISLVBCWLLVIDRNLQDDSHRSWHEHIFRMRQEPYLVKLVQTLSPQIILANIS
QIMPDAVFSTDVGOHQMAAQFVKTLPRMWSGSGTGWGCLPAIAGKILYPSPV
VCITGDSFQNMNIOELGTIAQYKLIKIIINNKQGMVROQCAFYGRYSHSMED
GAPNFVALKRSFGIDGOSISTROEMDSLNTIITKXGPMVIDCKVIEDNCVPMVAPG
KSNAMQMGIDKSNNEIKIKE"

ORIGIN

Query Match 32.6%; Score 638.4; DB 8; Length 2338;
 Best Local Similarity 61.8%; Pred. No. 2e-175;
 Matches 1050; Conservative 0; Mismatches 646; Indels 4; Gaps 2;

127 CAAGGCAACCGGGGCTTATATCCGATGAGTCCCTGAAAGCCCATGGGGTCAACAC 186
 571 CAAGAACAAACCGGGGCTTATATCCGATGAGTCCCTGAAAGCCCATGGGGTCAACAT 630
 187 ATTTTGGCTATCCCGGGGGAATTTTGGCTATATGATGATGATGATGATGATGATGAT 246
 631 ATTTTGGCTATCCCGGGGGAATTTTGGCTATATGATGATGATGATGATGATGATGAT 630
 247 GCGGCGGGGGAATTTGAGCATATTTTGGTGGCCATGAAACAGAGGCTTCCCATGGCG 306
 691 GAGCAAGGATATATTAAGCATATATGCTTGAACACAGAGGCTTCCCATGGCG 750
 307 GATGGTATGCGACAGCAGCAGTAAAGTGGAGTTTGTTCGGTACATGACAGCAGG 366
 751 GATGGTATGCGACAGCAGCAGTAAAGTGGAGTTTGTTCGGTACATGACAGCAGG 810
 367 GCGACTAATCTGGTACCGGCAATGGCCATGCCCATTGGACTGGTGGCCATGGTGGT 426
 811 GCGACTAATCTGGTACCGGCAATGGCCATGCCCATTGGACTGGTGGCCATGGTGGT 870
 427 ATTAAGGAGAGTGGGCGGCTGCGCATGATGATGATGATGATGATGATGATGATGAT 486
 871 ATTAAGGAGAGTGGGCGGCTGCGCATGATGATGATGATGATGATGATGATGATGAT 930
 487 TTTGGCATCACCTTACCGCATGTTAAGCATCTCTATGATGATGATGATGATGATGAT 546
 931 TTTGGCATCACCTTACCGCATGTTAAGCATCTCTATGATGATGATGATGATGATGAT 990
 547 GCTGCATTTTATCTAGAGCTTTCATCTGACACCGGCTGCTCCCGGGCGGTTTG 606
 991 CTTAAATATGTTGCTGAGGCTTCTCTATGATGATGATGATGATGATGATGATGAT 1050
 607 ATTCATTTTCCCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 666
 1051 ATTCATTTTCCCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1110
 667 GACCTTATCTACCGGCTTATGCGCCACAGCTTAAAGTAAATCCCGCAAAATTAAGCG 726
 1111 GAAAGCAAAATTAAGCAAGATTTGCGCATGATGATGATGATGATGATGATGATGAT 1170
 727 GCATTCATTTGTTGAGACAGCAGCAAAATCCCTTGTCTACGATGAGGAGGAGGCGATC 786
 1171 TCAATTCATTTGTTGAGACAGCAGCAAAATCCCTTGTCTACGATGAGGAGGAGGCGATC 1230
 787 GCGGCAATGCGCATGCGCAGGTCAGGAATTTGCGGAAGCTTCCGTTGCGGATACA 846
 1231 ACTTCGAGGCTCATGACTTATATTAATTAATTAATTAATTAATTAATTAATTAATTA 1290
 847 ACCACCGTATGAGGATTTGGGCTTTGACGAAACATCCCTTGGTGGGATATG 906
 1291 ACTCATTAATGAGGATTTGGGCTTTGACGAAACATCCCTTGGTGGGATATG 1350
 907 GGTATGATGCGCACCGCTATGCAACTTTCCTGACGCAAGTGAATTTGTTGATGCA 966
 1351 GGTATGATGCGCACCGCTATGCAACTTTCCTGACGCAAGTGAATTTGTTGATGCA 1410
 967 GTGGGGGCGGCTTTCGACGACCGGTTAATGCAAACTGACGAAATTTGCTAGCGCGCC 1026
 1411 TTGAGAGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1470
 1027 AAAGTATTCATTCATGACATGACACCGGCGGAGTGGGAAAAAACAAGGCTCCGATG 1086
 1471 AAAGTATTCATTCATGACATGACACCGGCGGAGTGGGAAAAAACAAGGCTCCGATG 1530
 1087 CCCATTTGGGGGAGTACGCGCATGTTTAAACAAGCTTTGACGCGGCGCGGAGATTG 1146
 1531 CCCATTTGGGGGAGTACGCGCATGTTTAAACAAGCTTTGACGCGGCGCGGAGATTG 1587
 1147 GATTACCCACCCATCCCATACACCAAGGATGTTAATTCGATGATCATTTGGCGG 1206

1588 AACTTGCAATTAGAGATAGACCATTTGGTTCATGAGCATGAGCGTTATATTTAGGTGGGCG 1647
 1207 ACCGATTAACCCCTCCAGGTCGCCCATATGAGATATCTATTTGCCCCCAGAGAGTACTA 1266
 1648 CAAGATATCCATTAATTTAGTACCTTAATTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1707
 1267 CAAGAAATTTGGTGGCGAGCCCGGAGTCTTACACACCGATGTTGGGCAACACCA 1326
 1708 GCGAATATTTAGCAATTAATGACATGATGATGATGATGATGATGATGATGATGATGAT 1767
 1327 ATGTGGGCGGCGGCTTTTGTGAAACATGAGCCCGGCGGATGATTTCCAGTGGCTTG 1386
 1768 ATGTGGGCGGCGGCTTTTGTGAAACATGAGCCCGGCGGATGATTTCCAGTGGCTTG 1827
 1387 GGTACGATGAGGCTTTGTTTACCTGCGCCATGAGAGCCAAAGTGGAG-GTGGGGAGCA 1445
 1828 GGCACAAATGGGATGATGAGGCTTACCTGACATGAGCGGCAAAATGATATCTGAAATCA 1887
 1446 GGGGCTCATTTGATGATGAGGAGATGCGGAGCTTCCAAATGATCTTCAGGAATGAGGAAC 1505
 1888 CTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1947
 1506 CTAGCCCATGACATGACATGACATGACATGACATGACATGACATGACATGACATGACAT 1565
 1948 ATAGCCCAATATATATATATGACATGACATGACATGACATGACATGACATGACATGACAT 2007
 1566 GTGCGTACGTGACATGACATGACATGACATGACATGACATGACATGACATGACATGACAT 1625
 2008 GTAGACAAATGACAAATGACAAATGACAAATGACAAATGACAAATGACAAATGACAAAT 2067
 1626 GGCATGACGACATGACATGACATGACATGACATGACATGACATGACATGACATGACAT 1685
 2068 GAGACATGACATGACATGACATGACATGACATGACATGACATGACATGACATGACAT 2127
 1686 AAGCGGAAATTTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1745
 2128 ACTAGGCAAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2187
 1746 ATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1805
 2188 ATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2247
 1806 AATGCGCAATGCTAGGTTT 1825
 2248 AAGCACAGATGATGCTT 2267

RESULT 9
 BX572098/c 346683 bp DNA circular BCT 23-SEP-2003
 LOCUS
 DEFINITION
 Prochlorococcus marinus MIT9313 complete genome; segment 4/7.
 ACCESSION
 BX572098.1 GI:33640689
 VERSION
 complete genome.
 KEYWORDS
 Prochlorococcus marinus str. MIT 9313
 Prochlorococcus marinus str. MIT 9313
 Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococaceae;
 Prochlorococcus.
 ORGANISM
 1 (bases 1 to 346683)
 Roca, G., Larimer, F. W., Lamerdin, J., Malfatti, S., Chain, P.,
 Ahlgren, N. A., Arellano, A., Coleman, M., Hauser, L., Hess, W. R.,
 Johnson, Z. I., Land, M., Lindell, D., Post, A. F., Regula, M., Shah, M.,
 Shaw, S. L., Steglich, C., Sullivan, M. B., Ting, C. S., Tolonen, A.,
 Webb, E. A., Zinser, E. R. and Chisholm, S. W.
 Genome divergence in two Prochlorococcus ecotypes reflects oceanic
 niche differentiation
 Nature 424 (6952), 1042-1047 (2003)
 JOURNAL
 MEDLINE
 22825698
 PUBLISHED
 2 (bases 1 to 346683)
 REFERENCES
 Larimer, F. and Roca, G.
 Prochlorococcus genome consortium

TITLE		Direct Submission
JOURNAL		Submitted (03-JUL-2003) Submitted on behalf of the Prochlorococcus genome consortium, the DOE Joint Genome Institute, Production Genomics Facility, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA, and the Genome Analysis Group, Oak Ridge National Laboratory, 1060 Commerce Park Drive, Oak Ridge, TN 37831, USA, laimer@ornl.gov
FEATURES	source	Location/Qualifiers
		1. .346683
gene		/organism="Prochlorococcus marinus str. MIT 9313"
		/mol_type="genomic DNA"
		/strain="MIT9313"
		/db_xref="taxon:74547"
		/complement(347. .637)
		/locus_tag="PMT0970"
		/complement(347. .637)
		/locus_tag="PMT0970"
		/codon_start=1
		/evidence=not experimental
CDS		/transl_table=11
		/product="possible Herpesvirus VP23 like capsid protein"
		/protein_id="CAE21145.1"
		/db_xref="GI:33640690"
		/translation="MKIKHSIRHIFPEGTELLCLRKVKOTLLSCDLYDDIPLLA
		WAGPAVDLIFSNNKRPDILRLGATFIVLALSSHVSPDVCASLMQHLS"
		620. .1015
		/locus_tag="PMT0971"
		620. .1015
		/locus_tag="PMT0971"
gene		/codon_start=1
		/evidence=not experimental
		/transl_table=11
		/product="possible hZIP transcription factor"
		/protein_id="CAE21146.1"
		/db_xref="GI:33640691"
		/translation="MDPSQKPYTSRSGDYRRPAVLATFLVRSIILMACCRGHE
		INDVSRPEVNGFPFGYTIIPDRVVVPSDAVYAAQERQNAQVARILEARAYLSV
		VDLEKQIAELQPSQEPVPDGLAAKT"
		1478. .1699
CDS		/locus_tag="PMT0972"
		1478. .1699
		/locus_tag="PMT0972"
		/codon_start=1
		/evidence=not experimental
		/transl_table=11
		/product="possible DNA polymerase (viral) N-terminal dom"
		/protein_id="CAE21147.1"
		/db_xref="GI:33640692"
		/translation="WPNQRLRALILFLKRVILAAQEEFFTPWQSKAGIDCIKRAM
gene		LSDVPEALADDQQLKQVLDOARWLVN"
		2436. .2690
		/locus_tag="PMT0973"
		2436. .2690
		/locus_tag="PMT0973"
		/codon_start=1
		/evidence=not experimental
		/transl_table=11
		/product="possible Vinculin family"
		/protein_id="CAE21148.1"
CDS		/db_xref="GI:33640693"
		/translation="MGPFMYGDERWILIRANALFECFRLWGLTNGILSKRKDLQVEKRP
		PKEVGNPKPRDLSFISRDDEFKRLQTLAESPSQED"
		/complement(2865. .3230)
		/locus_tag="PMT0974"
		/locus_tag="PMT0974"
		/codon_start=1
		/evidence=not experimental
		/transl_table=11
		/product="conserved hypothetical protein"
gene		/protein_id="CAE21149.1"
		/db_xref="GI:33640694"
		/translation="MMQDRNVLAGEPLEICSDPRTGYWRDQKCRITDGHDLGCHTVCV
		/complement(2865. .3230)
		/locus_tag="PMT0975"
		/locus_tag="PMT0975"
		/codon_start=1
		/evidence=not experimental
		/transl_table=11
		/product="possible Ribosomal protein L36"
CDS		/protein_id="CAE21150.1"
		/db_xref="GI:33640695"
		/translation="MRKPIIDPIFKSALAGILLCTSTISAQKRVIEGTYSYNCR
		EGLMKSRITIGRVNDKEMWITIERLEKISGRQVNVNDIKKSCQSSILISKARFAG
		DCELASDSFDSNMQIDVAFRFAHGNKHIFAIRYDGSRRCLSKNNPRDLPIET
		PGRSFLHNIERQADSSIIIEYHPNGMGYEVKKYKRLDSDPERSPSVDSMTQR
		"
		/complement(3918. .4007)
		/locus_tag="PMT0975"
		/note="Signal predicted by signalp 2.0 HMM (signal peptide probability 1.000) with cleavage site probability 0.694 at residue 30"
sig_peptide		4938. .5267
		/locus_tag="PMT0976"
		4938. .5267
		/locus_tag="PMT0976"
		/codon_start=1
		/evidence=not experimental
		/transl_table=11
		/product="possible Protein of unknown function DUF42"
		/protein_id="CAE21151.1"
		/db_xref="GI:33640696"
gene		/translation="MGFQALQVRGLNAGIDAPLPIAVGNSVFFIDALVANNFSP
		YDNYSIIINTVPMKRLCPGKKEGTLMLPGFYKGVSGMHELLEDQSRAYVQLISAG
		PAVLND"
		/complement(5257. .7404)
		/locus_tag="PMT0977"
		/locus_tag="PMT0977"
		/complement(5257. .7404)
		/locus_tag="PMT0977"
		/codon_start=1
		/evidence=not experimental
CDS		/transl_table=11
		/product="ABC transporter, multidrug efflux family"
		/protein_id="CAE21152.1"
		/db_xref="GI:33640697"
		/translation="WTLNPGREALIQARALEELLFSYTPENEGSITEILLGFCNAKL
		GRRQKRLQPLSGASIKELNENDIHRKLEAPLDNQAEYPLVAFVDEHNPFALYR
		HKGQWMLFPDQESHRLCDNSLKLALAEVVPSPLEVLGLAVLRFAPNTEPLAT
		ALVITSAVVMFLTSLIPMLTVLVRILPIONQNTLFGQLTAVVILVIGSMATEFQS
		LMILRSTADRLQSAVNDRLMLPMSFISKTYTGLASRVASISOLROLIGNVTS
		TLISLPAISPFVIMFVYDQALMAASATVIAACILMITRISIQLOEPILSEGAET
misc_feature		TNPSLOAVMGQINSAGSEPYLVRMLREIRVASIQRSNVISQDMEQYGLVPL
		ASLPMFAVAVQVINSPEILNQLVAFISFNAFSGNGALTGAINILAVAVARAA
		VLMRRAEPLVADVERKGYQPEAIVHSLNGEFLRGLISYEFPSSEQLFNSLTLLT
		QHTAITGASGCGKTLVLMGLGISIPQTEGLVLDGIPLTQLAIRAYRQLGVNQTAR
		LNTGSIYEIVICGVGSRSEDEVMQALEQALADEVAMPQLETILSDSGNLSGGQVO
		RIALRALITOPKVLIMDEATSLDNRSERITOTIKELGMTRISIAHRLSTIQADQ
		YILERNQKXEMSGKWEIKSHGLASMLASH"
		/order(complement(6868. .6933), complement(6757. .6822),
		complement(6523. .6588), complement(6451. .6507))
		/complement(6187. .6252), complement(6076. .6141))
gene		/locus_tag="PMT0977"
		/note="6 probable transmembrane helices predicted by TMHMM2.0"
		/complement(7401. .9617)
		/locus_tag="PMT0978"
		/locus_tag="PMT0978"
		/complement(7401. .9617)
		/locus_tag="PMT0978"
		/codon_start=1
		/evidence=not experimental
		/transl_table=11
CDS		/product="ABC transporter , multidrug efflux family"
		/protein_id="CAE21153.1"

D	278847	CTCGACGAGTTTTCACCGAATCCCGGCTCCAGAACTTTCCTCAATTGAGCGGCTCTCGAT	278788
Q	736	TTGTTGAGACAGCCAGAAATCCCTTGCTTACAGAGGGGAGGGGCGATGCGCGCAAT	795
D	278787	TTAATTGAAGATGCTCAAGAGGCTCTCTCTTACGTGCGGAGGAGTTATTTCTGCCGG	278722
Q	796	GCCCATGCCCAAGTGCAGAAATTTGCGGAAAGGTTCCAGTTGCCGTAACACACCCTCG	855
D	278727	GCTCAGAGAGTCTTAAGSGTTTTCAGAAAGTCCACGATACCTGTCACACACCAATWG	278668
Q	856	ATGGAATTTGGGGCTTTTACGAAAAACATCCCTTGGGAGGTATGTTGGGATATCAT	915
D	278667	ATGGGTAAGGGGGCGTTTGATATAGGCAATCTTTGGCTTTGGGAGTGTGGGAAATGAT	278608
Q	916	GGCACCGCTATAGCAACTTTCGCTTCACGATGTGATTTGTTGATTTGCAGTGGGGCC	975
D	278607	GGACATGCTTAAGCAACTTTGCAATTTCTGAATGTATTTGTTGATGCTGTTGGTCA	278544
Q	976	CGTTTGCAGACCGGGTAACCTGGCAACTAGACGAATTTGCTAACCGCGCCAAATATTT	1035
D	278547	CGTTTGAGTAGCCGTGTATTCAGGCAAGCTTGTACTTTTGGCCCCGCAAGATGATTT	278488
Q	1036	CACATTCATCGAACCCCGCGAGGTGGGAAAAACAAGGCTCCCGATGTGCCCATTTGTG	1095
D	278487	CATTTGAAATTTGATCCAGACAGATAGTAAAGAACGACGTCTGATGTTGCCGTAATG	278422
Q	1096	GGGAGATGACGCCATGTTTAAACAAGCTTTTGCACGGCGCCCGGAATTGGAATTAACC	1155
D	278427	GGTAGATGTTGGTGCAGTGTGTGAAGTTATTCGACTTGAAGTAAAGCAAAATGCT---	278372
Q	1156	ACCATCCCATATCCACCCGAGCATGTTAAATGSCATGTATCAATTGGCGAGCATTTAC	1215
D	278371	-----TGAATTGAGACTTCAAGCTGGCTACTACCTGATTTGAGAGTTGGAABACCTTTAC	278311
Q	1216	CCCCCTCAGGTGCCCATATAGGATATTTGCCCCCGACGAGTATGTAACAGAAATT	1275
D	278316	CCCTTGATGATCCCTCGCTAGGAGGGGGCCAAATTTATCTCCAGAGGTGCTTTGGCGATA	278257
Q	1276	GGTGGCAGGCCCCGAGTGCCTACACACAACGATGGGACAAACCAATATGGGCG	1335
D	278256	AGGAGCTTGCTCTTGAAGCTTACATACACTGATGTGGGCGACAGATCAATGTGGCG	278197
Q	1336	GCCAGTTTTTGAACAATGGCCCCGCGAGTGAATTTCCAGTGTGCTTGGGTACGATG	1395
D	278196	GCTCAATATTACGAAATGGGCTGCTGTCAGTGAATCAATAGGCGGGATTTGGGAACGATG	278137
Q	1396	GGCTTTGGTTTACCTGCGCGCATGGGAGCAAGGTGGAGGGGAGAGC-AGTCATT	1454
D	278136	GGTTGCGGATGCGACATCTATGTGTTCAGTTGCTTTGCCAGATGAACAGTTGTT	278077
Q	1455	TGCATCAGTGAATCCGACTTCCAAATGAATCTTCAGAACTGGGAACCTTACCCAG	1514
D	278076	TGTAATGCTGTGATTTCCAGCATCTGATGAATATTCCAGAGTTGGGAACACTGATTCAG	278017
Q	1515	TAGCAATCCAGGTTAAACTATATNTCCATTAACGTTGGCAGGGGATGTCGTCAG	1574
D	278016	TACAACTCCACACAGATTGTGATTTGAACAACACTGGCAAGGGATGTTCTGTCAG	277955
Q	1575	TGGCAACAACTTTTACAGAGAAGTTATTTGCTTTAAACATGTCCACAGGCAATGCA	1634
D	277956	TGGCAGGAAGCTTTACGAAAGTGTATTTCAGCCACATGACATCTTCCGGGACATGCTT	277897
Q	1635	GACATTAATTCCTCTGTGAAGCTTATGGCATCAAGGATTTACGTGTCCGACAGCGGAA	1694
D	277896	GATTTGCTGCAATGGCAGAGGCTTTTAAAGTTGCGGGGTGTTGATCATCTGAGGAAAG	277833
Q	1695	GATTTGGCCCCGCGCATCCCAAAATGTAGCCCAATAGTCTGCTGTGATGTGATGTG	1754
D	277836	GATCTCAGGTCAAGTTTGAAGACGCGCTTGGCATCCAGGTCCGATGCTCATGACGTG	277777
Q	1755	GTCGTCAAAAAAGATGAATCTTATACCTATGATTTGCCCGGATATGTAATGCCCAA	1814

Db 277776 CATGTGCACGATGATAAATGCTATCCGATGTTCTCCTGGCAGAGCAGCTCAA 2777717
QY 1815 ATGCTAGGTTTACCGGAAGTCCGGTAC 1842
Db 277716 ATGGTGGTTTATCTAACCCCTGAAC 277689
RESULT 10
PTCEPTR/c 4099 bp DNA linear PLN 26-OCT-1993
DEFINITION Porphyra umbilicalis chloroplast transfer RNA-Tyr, transfer
RNA-Thr, acetylglutamate synthase, N-acetylglutamate kinase, transfer
RNA-Met, transfer RNA-Ala, transfer RNA-Ser, and transfer RNA-Asp
genes, complete cds.
ACCESSION M94625.1 GI:3431329
VERSION M94625
KEYWORDS N-acetylglutamate kinase; acetylglutamate synthase; transfer RNA-Ala;
transfer RNA-Asp; transfer RNA-Ser; transfer RNA-Thr; transfer
RNA-Tyr; transfer RNA-Met.
SOURCE Chloroplast Porphyra umbilicalis (laver)
ORGANISM Porphyra umbilicalis
Eukaryota; Rhodophyta; Bangiophyceae; Bangiales; Bangiaceae;
Porphyra.
REFERENCE 1 (bases 1 to 4099)
AUTHORS Reith,M. and Muniholland,J.
TITLE Two amino-acid biosynthetic genes are encoded on the plastid genome
of the red alga Porphyra umbilicalis
JOURNAL Curr. Genet. 23 (1), 59-65 (1993)
MEDLINE 93153832
PUBMED 8381336
COMMENT Original source text: Chloroplast Porphyra umbilicalis (strain
Avonport, organelle Chloroplast Porphyra umbilicalis) DNA.
FEATURES
source location/Qualifiers
1..4099
/organism="Porphyra umbilicalis"
/organelle="plastid:chloroplast"
/mol_type="unassigned DNA"
/db_xref="taxon:2786"
105..185
/gene="trnY (GUA) "
105..185
/gene="trnY (GUA) "
/product="trnY (GUA) "
/note="putative"
/citation=[1]
/anticodon=(pos:139..141,aa:Tyr)
198..269
/gene="trnT (GCU) "
198..269
/gene="trnT (GCU) "
/product="trnT (GCU) "
/note="putative"
/citation=[1]
/anticodon=(pos:230..232,aa:Thr)
complement(275..2047)
/gene="11bB"
complement(275..2047)
/gene="11bB"
/EC_number="4.1.3.18"
/note="putative"
/citation=[1]
/codon_start=1
/product="acetylglutamate synthase"
/protein_id="AA03052.1"
/db_xref="GI:343130"
/translation="MLSKQIIIGSEKTRPALDISVHGVHIFGYGAILPIYDEL
YAMELSLIIKNIIVRHEQGSADAYSRSKGVCFATISGPATMLVSGIATAND
SVPIALITGVGRPFITGDARQEVDFIQTILPIVHSGVYADPDMRIVAEAFYICK
HGRGPVLIDVQKVGLEKFNYSVSPQVXIPGCRPLSNLKSQIIMAAKMOOSSO
PLVITGGGAILISDAHSIIKELVLDYKIPVTTLLGKGFNDSDFCIGMLGMGTAY
NFAVSECDLIIAIGARFDDRYTGKLDERCAQVYIHVDIPABYGRKRIQVAVGV
TEVTSLSNLKNNFKYPPDIIISWGRIHNRQCYLVPKSTSTISPEIILTTNQ

CVSGDSSFGMMNDELGTIIAOKYKLPKIVIIINRMQGMVPMOQAFYGERYSHRMTG
ARRFOKLARFGCIKATVANNRQMMESSLKDKANKYRPVILLDCQVTENENCIYMAVAPGK
SNQWMTGIAKPPQGTASNYVSRNI"
gene 2485..3336
/gene="argB"
2485..3336
/gene="argB"
/EC_number="2.7.2.8"
/note="putative"
/citation=[1]
/codon_start=1
/product="N-acetylglutamate kinase"
/protein_id="AA03053.1"
/db_xref="GI:343131"
/translation="MLTNERVRYVLSDYTIILQESSRIIVIKYGAAMNOKLKQHI
SDIVFISFEIGRLPILVHGGPEINFMDDLKILPRENGVRVTDQPTMDYEMVYGR
VNDVLVASINKQGSXVLSGKDLITRPSDKENLGVGVQVNDVTLLEILINN
YLPVLASVADKQOQSYINADTVAGEIAPARNASKLILITPGLIRNASADATLIS
HLSIQARDLTITAVISGGMIPKVNCCIRSLAQVASAHIIDGRIDHALLILTDQG
IGSMLVV"
3504..3577
/gene="trnFM(CAU) "
3504..3577
/gene="trnFM(CAU) "
/product="trnFM(CAU) "
/note="putative"
/citation=[1]
/anticodon=(pos:3538..3540,aa:Met)
3583..3654
/gene="trnA(GGC) "
3583..3654
/gene="trnA(GGC) "
/product="trnA-Ala"
/note="putative"
/citation=[1]
/anticodon=(pos:3616..3618,aa:Ala)
3699..3788
/gene="trnS(GCU) "
3699..3788
/gene="trnS(GCU) "
/product="trnS(GCU) "
/note="putative"
/citation=[1]
/anticodon=(pos:3734..3736,aa:Ser)
3928..3992
/gene="trnD(GUC) "
3928..3992
/gene="trnD(GUC) "
/product="trnD(GUC) "
/note="putative"
/citation=[1]
/anticodon=(pos:3962..3964,aa:Asp)
ORIGIN
Query Match 30.8%; Score 603.6; DB 8; Length 4099;
Best Local Similarity 60.5%; Pred. No.3,5e-165;
Matches 1029; Conservative 0; Mismatches 669; Indels 4; Gaps 2;
QY 132 GCAAAACCGGGCTTATATCCTGATGATAGCTGAAACGCCATGGGGTCAACACATTIT 191
Db 2018 GAAAACTGCGTCGTTGCTCTTTTAGATGCAATGTTGTCACGGAGTTATTCATATT 1959
QY 192 TGGCTATCCCGGGGGCAATTTTGGCCATCTATATGATGACTGACCGCTTTGAAGCGGC 251
Db 1958 TGGTATCCCTGGGGGCTATTTTACCAATTTATATGATGAACTTTATGCTTGGAGAGACT 1899
QY 252 GGGGGAATTTGACATATTTTGGTGCCCATGAACAGAGCTTCCCATGCGGGCGATGG 311
Db 1898 TTCATTATCAAGAAATATTCTTGTAAGACATGAACAAGGTGCTTACATGCTGCAGATGC 1839
QY 312 GTATGCAGAGCCACAGGTAAAGTGGAGATTGTTGCTATCATCTGCAGCCAGGGCCAGC 371

QY 372 TAACTGTGACCGGCGATTGGCAATTGCCATTGGATTCGGTGGCCATGGTGTGATTAC 431
 Db 1778 TAAATTGATATCCGCGATGACTGCGCAATTGACTCGGTCTCTATTACTATAC 1719
 QY 432 TGGAGAGTGGCGCGTCATGATTGGTAGCGATGCTTTCAGAGAAATTGACATTTTGG 491
 Db 1718 TGGCAAGTGGGACGACCTTTTATAGGAACAGATGCAATTCCAGAGTTGATATTTTCG 1659
 QY 492 CATACCTTACCGATCGTTAAGCACTCTTATGTGTACGTAGTGTGGCGGATATGCTCG 551
 Db 1658 CATTAACATTAATCTTAAACATTAATGATGATGCGGACCTTAAAGAACATGCTGAG 1599
 QY 552 CATGTTACATGAGGCTTTCATCTTGTAGACACCGGTCGTCGCGCGGCTTTTGAATCG 611
 Db 1598 AATTGTTGCTGACCGTTTATCATTTGCAAAACATGTAGACCTGCGCCATTTTATTTGA 1539
 QY 612 TATTCCCAAGATGTGGCTTAAAGAAATGTAGTACATTCCTCGACCCCGGTGACGT 671
 Db 1538 TGTACCAAAAGATGTGGACTAGAAAATTTAATTATTTCTGTGTGAGCAGAGGCAAGT 1479
 QY 672 TAACTACCGGGGTTATGCGCCCAAGTTAAAGGTAATCCCGCAAAATTATGCGGCAAT 731
 Db 1478 TAAATCCCTGTGCTGACAGCTCTATCTAATCTAATAATCCCTGCAAAATCTAATGCTGC 1419
 QY 732 GCAATTGTGAGAGGCGCAAAATCCCTGTGCTGCTAGTAGGAGGAGGCGATGCGCCG 791
 Db 1418 AAAATGATTCAACAATCAAGTCAAGCATTTATTAATATTGTTGTGTGAGACTATTAATTTC 1359
 QY 792 CAATGCCATGCCAGGTGAGAGAAATTGGGAAAGGTTCCAGTTGCGGCTAACACAC 851
 Db 1358 TGAATGCTACATTAATTAAGAAATGATTGACTTATTAATAATACAGTAATCTACAC 1299
 QY 852 CCGTATGGGAATTGGGCTTTTACGAAAACATCCCTTCGTTGGGTATGTTGGGTAT 911
 Db 1298 TCTTATGGGCAAGGATTCCTTAAATGAGATGAGATTTTGTCTAGATGCTCGGAAT 1239
 QY 912 GCATGCGCACCGCTATGCACTTGGCGTCAGCGAATGTGATTTGTTATTCGATGGG 971
 Db 1238 GCATGGGACGTGCGTACCTTAATTTGAGTAGGATGTGATCTTCTGATTCCTCTAGG 1179
 QY 972 GGGCGGTTTGCAGCAGCGGGTAATGTCGCAACTAGACGATTTGCTAGCGCGCCAAAGT 1031
 Db 1178 AGCTAGATTGATATGATAGGTAAGTCTGTAATTAATTAATGATTTGCTGTATCTCAAGT 1119
 QY 1032 AATTCACTGATGATGACCCCGGCGAGGTGGGAAAAACAGGCTTCGCGATGTGCCAT 1091
 Db 1118 TATTCATGTAGATTTGATGATCCAGCGGAATGGAATTAATACCTCAAGTTGCTAT 1059
 QY 1092 TGTGGGGATGTAGCGCATGTTTGAAGACGTTTTCAGCGGCGCGGGAATTTGATTA 1151
 Db 1058 TGTGTGTGATGTACAGAGTTGTTACTTCACTTTTGA--TTTACTGAAAAATAATTT 1002
 QY 1152 CCCCACCCATCCCATACCAACCCAGGATGTTAATGCGATTGATCATTTGGGACGA 1211
 Db 1001 TAAGCTTATTCAGAGCAAAATTATATCTTGGCAAGAAATACATGCTGGCGTACGA 942
 QY 1212 TTACCCCTCCAGGTGCGCCCATATGAGATCTATGCCCCCAGAGAGTATACAGA 1271
 Db 941 ATACCTCTTTTATACCAAAAAATCACTAGCATTTTCAACAAGAAATACTTGTATC 882
 QY 1272 AATTGTGCGCAGGCGCCCGATGCTTACACACCGCATGTGGGACAAACCAAAATGTG 1331
 Db 881 AAAAAATCAATTACGCAAGATGCTTATTTTACTACGATGTGCGCAGCATCAAAATGTG 822
 QY 1332 GGGCGGCGAAGTTTGAACAATGAGCCCGCGCATGATTTCCAGTGTGCTGGTGTAC 1391
 Db 821 GTAAGCTCAATTTCTTAAAGTGAATTAAGCAATGTGATTTCTAGTGTGAGCTAGAAAC 762
 QY 1392 GATGGGCTTGTGTTTACCTGCGCATGAGACCAAGTGGAGTGGGGGACGAGC-GGT 1450
 Db 761 AATGGGCTATGTTTGCACAGCTATTTGTTGCCCAAGTACACACCTTAACGAATTAAT 702

QY 1451 CATTGCATCAGTGGAGATGCGAGTTCACAAATGATCTCAGGAACTGGGACCTTAGC 1510
 Db 701 AATTGTGTAAGCGGTGATTAAGTTTCAATTAATTAATGCAAGAAATTAGAACTATTGC 642
 QY 1511 CCAGTACGACATCCAGGTTAAACTATTATTTCATTAACGTTGGAGGGGATGTGCG 1570
 Db 641 CCAGTAAATTAATCCATTAATAAATGTTATTAATTAATTAATGCTGGCAAGGCATGTAG 582
 QY 1571 TCAGTGGACAAACTTTCTAGAAAGACGTTATTCGCTCTTAACAATGTCGCCGCAAT 1630
 Db 581 ACAGTGGGACGAGATTTTACGGAAGAAATCTCATTCATTCAGAAATGACAGAGAAC 522
 QY 1631 GCCAGACATTAATCTCTCTGTGAGAACCTATGAGCATCAAGGGTATTACTGTGCGAAAGC 1690
 Db 521 TCTTATTTTCAAAAGCTGTGAGAGCTTTGTGATTAATTAAGCTTTACTGTAATTAATAG 462
 QY 1691 GGAATTTTGGCGCGGATGCGCGAATCTGACCCACATATGCTCTGTGTGATGGA 1750
 Db 461 ACAAATATATGAAATCTCTTTAAAGATGATTAATAATACCTGCTGCTCCGTTTATTA 402
 QY 1751 TGTGTGTGCTCAAAAAGATGAAACTGTATACCTATGATTTGCCCGGCATGATTAATGC 1810
 Db 401 TTGCCAATTAATCAAAATGAAATTTGTTATTCATGATGTTGACACAGGAAAAATATATGC 342
 QY 1811 CCAATGCTAGGTTTACCGGAA 1832
 Db 341 ACAAATGATAGGAAATAGCAAAA 320

RESULT 11
 PU38804
 LOCUS PU38804 191028 bp DNA circular JUN 27-MAR-1998
 DEFINITION Porphyra purpurea chloroplast, complete genome.
 ACCESSION U38804
 VERSION U38804.1 GI:127652
 KEYWORDS
 SOURCE chloroplast Porphyra purpurea
 ORGANISM Porphyra purpurea
 Eukaryota; Rhodophyta; Bangiophyceae; Bangiales; Bangiaceae; Porphyra.

REFERENCE 1 (bases 1 to 191028)
 AUTHORS Reich M.E. and Munkolund J.
 TITLE Complete nucleotide sequence of the Porphyra purpurea chloroplast genome
 JOURNAL Plant Mol. Biol. Rep. 13 (4), 333-335 (1995)
 DIRECT SUBMISSION
 SUBMITTED (17-OCT-1995) Michael E. Reich, Marine Biology Section, NRC Institute for Marine Biosciences, 1411 Oxford Street, Halifax, Nova Scotia B3H 3Z1, Canada

FEATURES
 source
 1..191028
 /organism="Porphyra purpurea"
 /organelle="plastid:chloroplast"
 /mol_type="genomic DNA"
 /strain="Avonport"
 /db_xref="taxon:2787"
 181..1053
 /gene="chlL"
 181..1053
 /gene="chlL"
 /codon_start=1
 /product="prochlorophyllide reductase iron-sulfur ATP-binding protein"
 /protein_id="AAC08073.1"
 /db_xref="GI:127653"
 /translation="MKLAAYGKGIKSTTSCNISVALSKRKKVLIQIGDPKHSIF TLKFLIPITIDTLOADYHEDVWPEDVYKGYGVDCVBAAGSPGAGCGYVVE TVKLKELNAFDEVDILPVLGAVVCGFAAPLNTYDCLITDNGFDALFANRFA ASVERKARTSLRAGLVGRTRKRDLDKVICVPMPEVLEPLIDIVSRVKGKT LFEABSDTNLAAYCDVYLNINADQLIRPBGVFKEAADNELSLSDPLNPKSNTT KSTADESELIMV"

QY	972	GGCCGCTTTGACGACCGGGGTAACTGGGAAACATGAGCAATTGCTGACGCGCCAAAGT	1091
Db	119082	AGCTGAATTTGATGATAGGGTAATCTGTAAATTAAGTAATTCGTTGTAATGCTCAAGT	11914
QY	1032	AATTCAATTGACATGACCCCGCGAGGTGGGAAAAAACAAGGCTCCGATGTGCCAT	1091
Db	119142	TATTTATGTAAGTAATGATCCAGCCGAAGTTGTGTAATAAGAAATACCTCAAGTTGCTAT	119201
QY	1092	TGTGGGGAGATGTAAGCGCATGTTTTTGAACAGCTTTTGACGGGGCCCGGAATTGATTA	1151
Db	119202	TGTTGTGTATGTTACAGAAAGTTGTACTTCACTTTTGA---TTTACTGAAAAATAATTT	119258
QY	1152	CCCCACCCATCCCATACCAACCCAGGACATGGTTAAATGCAATGATCATTTGGCGGACGA	1211
Db	119259	TAAAGCTTTTCCAGAGCAAAATTATCTTTGGAGAAAGAAATACATGCTGGCGCTACGA	119318
QY	1212	TTAACCCCTTCACAGTGTGCCCACTATAGAGATCTATTTGCCCCCGAGAGGTATCACGA	1271
Db	119319	ATATCCCTCTTTAGTACCAAAAAATCAACTAGCATTTTCACACAAATACTCTGTAC	119378
QY	1272	AATTGGTCGACAGGCCCCCGATGCTCTACACCCAGCATGTGGGACAAACCAATGTG	1331
Db	119379	MACAATTCATTTAGCGCAAGATGCTTATTTTATCTATGATGTGGCCAGATCAAAATGTG	119438
QY	1332	GGGGGCCAGTTTTTGAACAATAGGCCCGCCGATGTATTTCAAGTCTGCGTTGGGTAC	1391
Db	119439	GTCAAGCTCAATTTCTTAAAGTGAATCTAAGCATTTGATCTGTGACTGACATAGGAAC	119498
QY	1392	GATGGGCTTGTGTTTAACCTGCGGCATAGGAGCCAAAGTGGAGATGGGGACGACG-CGT	1450
Db	119499	AAATGGCTATGTTGTTCGACAGCTATTTGTGCCAAGTACACACCTTAACGAATAGT	119588
QY	1451	CATTGTCATCAGTGAAGATGCGACGCTTCCAAATGATCTTCAGAACTGGGAAACCTTAGC	1510
Db	119599	AATTTGTGTAAGGGGATTCAGGTTTCAATGATTAATGCAAAATTAAGAACTATTTGC	119618
QY	1511	CCAGTACGACATCCAGGTTAAACATTAATTTCTCAATACGCTTTGGCAGGGGATGTGCG	1570
Db	119619	CCAGTAAATTAATACCTAATAAAATGTTATTAATAAATGCTGGCAAGCATAGGTTAG	119678
QY	1571	TCAGTGGCAACAACTTTCTACGAAGAAGTTATCTGCTCTCAACATGTCGCCAGGCGAT	1630
Db	119679	ACAGTGGCAGCAAGCATTTTACGGAGAAAGATCTCTCATTTACAGAAATGACAGAAAGAGC	119738
QY	1631	GCCAGACATTAATCTCTCTGTGAAGCCTATGGCATCAAGGTAATTACTGTGGCAAGCG	1690
Db	119739	TCTTAATTTTCAAAAGCTGTGGAAGCTTTTGTGTAATTAAGCTTTTACTGTAATAATAG	119798
QY	1691	GGAAGATTGGCCCGCGGATGCGCGGAATGCTAGCCACAATGCTGTCTGTGGGATGGA	1750
Db	119799	ACAAAATTAAGGAATCTCTTTAANAAGATGCTATGAATATCCCTGTGCTCCGTTTATTAGA	119858
QY	1751	TGTGTGTGTCAAAAAAGATGAAGAATGTTATCCCTATGATATGGCCCCCGCATGAGTAATGC	1810
Db	119859	TTGCCAAGTAATCGAAAATGAANAATTTGTTATCCAATGGTTTCACAGAGAAAAGTATGC	119918
QY	1811	CCAAATGCTAGGTTTATCCGGAA	1832
Db	119919	ACAAATGATAGGAATAGCAAAA	119940
RESULT 12			
AB002583/c			
LOCUS	AB002583	149987 bp	DNA
DEFINITION	Cyanidioschyzon merolae chloroplast DNA, complete genome,		
ACCESSION	AB002583	strain:10D.	
VERSION	AB002583.1	GI:30409147	
KEYWORDS	chloroplast Cyanidioschyzon merolae		
SOURCE	Cyanidioschyzon merolae		
ORGANISM	Eukaryota; Rhodophyta; Bangiophyceae; Cyanidioschyzon.		
REFERENCE	1		

AUTHORS Ohta,N., Matsuzaki,M., Misumi,O., Miyagishima,S., Nozaki,H., Tanaka,K., Shin-I,T., Kohara,Y. and Kuroiwa,T.
TITLE Complete Sequence and Analysis of the Plastid Genome of the Unicellular Red Alga *Cyanidioschyzon merolae*
JOURNAL DNA Res. 10, 67-77 (2003)
REFERENCE 2 (bases 1 to 149987)
AUTHORS Ohta,N., Matsuzaki,M., Tanaka,K., Shin-I,T., Kohara,Y. and Kuroiwa,T.
TITLE Direct Submission
JOURNAL Submitted (02-Apr-1997) Niji Ohta, Saitama University, Department of Molecular Biology and Biochemistry; 255 Shimo-ohkubo, Saitama, Saitama 338-8570, Japan (E-mail:niji@molbiol.saitama-u.ac.jp, Tel:81-48-858-3848, Fax:81-48-858-3384)
FEATURES
source
1..149987
/organism="Cyanidioschyzon merolae"
/organelle="plastid:chloroplast"
/mol_type="genomic DNA"
/strain="10D"
/db_xref="taxon:45157"
1..1398
/gene="ycf24"
1..1398
/gene="ycf24"
/codon_start=1
/transl_table=1
/product="ABC transporter for iron-sulfur cluster formation (sucB)"
/protein_id="BAC76097.1"
/db_xref="GI:30409148"
/translation="MNQQLNLDPMNREYAYGFSPTSDYEPFPIGLNEQIVKMSAKCEPQMDLFRKAVHKKOMRQPMARVARETINFEOLLYAAPTDEDVQSLATLGVKSSKVALDVPDVSISGTGFHQAQKDAVITCSSEAIQKPSIYKXKYLAVYI GDNFSAIANSVPSDGSFCYTPDTACMEISTFRLOKSGSGQFERTLITAEAGASVSYLSGCTARVDRHQLHAAVELIALDHASIKYSTVQNWVSQDQEGKGIYNFVTKG LCAEGAKTISQVETGSAITWKTSPSCLANNTRADYSVALTDGYQOQADTGS KMHLGANSKSR1ISKGIQCGKSNMTYRGQVMTSAQAKAMNSQCSLLIGKATANTYPIEIANVSARVHEATTTQIGEEQFYVQGRGINQDAVNWGLGFCQGVFLQPMERA TEVVELTKLEGNCC"
1399..2139
/gene="ycf16"
1399..2139
/gene="ycf16"
/codon_start=1
/transl_table=1
/product="ATP-binding subunit of ABC transporter for iron-sulfur cluster formation (sucC)"
/protein_id="BAC76098.1"
/db_xref="GI:30409149"
/translation="MLKIEQLKGIKDRLLIEGIDAHIEAGSVHAIMGPNQSGKSTLAKVIAHPAYKILIEGFWQDESINDKAEAAQAGIFAFQNPPLTIPGVNMDLIMPA YKTKHOSIDAISFLEYIMPKLELVGADASPLHRYVNDGSGEKKKEITLQALIEP RLALIDEIDSGLDLDAKQIASHLMLKTTQNAVYLTHPRALLEYIKPDYVHWQGR RIIISGDASIAKYEIRYGYDEVKKAAV"
complement (2171..2275)
/gene="ycf12"
complement (2171..2275)
/gene="ycf12"
/codon_start=1
/transl_table=1
/product="unknown"
/protein_id="BAC76099.1"
/db_xref="GI:30409150"
/translation="MNHWVLAQLTVLAFLVAVGPITLWLSNKKDALI"
complement (2193..3065)
/gene="ycf62"
complement (2193..3065)
/gene="ycf62"
/codon_start=1
/transl_table=1
/product="cell-cycle protein mesJ family"
/protein_id="BAC76100.1"
/db_xref="GI:30409151"

/translation="MYTPMHNKRONRFQLANLTHSHSFSSHWLAISSGQSLCLLPKMDLVHMHVWVHFDHMSGSLSAARVYLSRMYHLMAYEFTALPITEQDARYRANILQYOTHSIATHTASDDETYLDDWISLRHPEGIWRLCHLSGSLPEPLHLTRAQTHWFTLVHYLPIWSDVSNQITPKRQHLQOLIPIKQINPREHTLSARLSARLSPLVYPHWMSLPIWLOQIATQWGLTSLQOLYINNKIELARSSINRFS FQDSRSYHFNAL"
3086..3157
/gene="trnK"
3086..3157
/gene="trnK"
/product="trnK-lys"
/note="codon recognized: AAA"
/anticodon=(pos:3118..3120,aa:lys)
3183..3914
/gene="trpA"
3183..3914
/gene="trpA"
/codon_start=1
/transl_table=1
/product="tryptophan synthase alpha chain"
/protein_id="BAC76101.1"
/db_xref="GI:30409152"
/translation="MLIAYLTAGADPINTTKEAVMKLAKGADVIEIGVPSDALADG AIIQKSKQALMNGFHLDMWLSVEVNEIEVPLIAYNQIMHYGVKWKXIVAH NVKGLIVPLPIYESKTLRQICDRYGANITWISPTNKTROELARACKWIVYISR TGYTGLETSPDKQIPKTLGELKKYTKAPIALGFGISKEQVRLVSKWADGVITGSAC MQLLEKGVQDLSWLSVWKSFCP"
3800..5140
/gene="infB"
3800..5140
/gene="infB"
/note="might be a partial remnant"
/codon_start=1
/transl_table=1
/product="initiation factor 2"
/protein_id="BAC76102.1"
/db_xref="GI:30409153"
/translation="MGGCGWCKKCMHAFIRKRSSTIRMDKDEKIIISVGTIYAD LAKQKPYTPERIRYAFQCGVIRLINDYDEIMQLQCKGEMQIYVEDTADAEV DYARDDEDEMEIIVASAVEKLEQAPITMSHWISLAIKPGVFLQMRKALANLQOII DKDVNHPRGPULIVESRKQNLGIIITLLIRKKSLLKVGDAVIEGNKIRAMQINQRI KEVAGEIVVSGFNQVSLGNTLYSAKKGEKASNFQIMQKICHPKGMIVASE SESALESVCFPLENKKQLLSAEGSIRKQVLAQTSGSILVAVQTCVMINMNVF PTVKIIKAKNFEDLLAQLESRYKVDRIVVSRLVQVPRQSGIAGCVVILQGEFKTGE VEIWRNQYQARARLSIKIKHQAKEYAVAPQEGIFKMDTWQVDEBIVLYN"
complement (5140..5212)
/gene="trnT"
complement (5140..5212)
/gene="trnT"
/product="trnA-Thr"
/note="codon recognized: ACA"
/anticodon=(pos:5177..5179,aa:thr)
5213..5785
/gene="rps4"
5213..5785
/gene="rps4"
/codon_start=1
/transl_table=1
/product="30S ribosomal protein S4"
/protein_id="BAC76103.1"
/db_xref="GI:30409154"
/translation="MSRYIGPVRVIRIRIGILPAFTNKS PNKRTGVGSEHAHAKRLS EYAVQLQERKQLQYVYGINQQLARVFRQAKRSRSTGIELKMLETRLDHVRAGF APTLPARQDLVNHGVKVGNGQVTLASFACQVNHILIEVAKSPSSAQLPPYLVQENQF VKMIOPEKWDAPRVNELLVVEYTRVGA"
complement (5788..6561)
/gene="ycf80"
complement (5788..6561)
/gene="ycf80"
/codon_start=1
/transl_table=1
/product="ORF257"
/protein_id="BAC76104.1"

```

/db_xref="GI:30409155"
/translation="MMDYIYEVRLTEDEPKMQLLQMFVSPVELEHLLLSROISV
CISEKPSFIDYQFSYFPKNTOLLVQNRASLEACPVIYVNPQOEILLSPRQWE
LEKAPPTLNMVFDHLEPRNPSTYLAFFHPDADHMYLKSIVNPFERKNSFVEGD
LFSFKLNTDILKIVLPLRPLRQCKMSKEFGTFVFWQMDHEMLFESKMDAQPFKN
QLSKTIYDVISLESVLESDLDQARCTVPLSLNOSL"
gene
complement (6539..6766)
/ gene="rp128"
complement (6539..6766)
CDS
complement (6539..6766)
/ gene="rp128"
complement (6539..6766)
/ codon_start=1
/ transl_table=11
/product="50S ribosomal protein L28"
/ protein_id="BAC76105.1"
/ db_xref="GI:30409156"
/ translation="MSRVCPSTLKCNHAEVTVYSHKRNHRQFVNLQKKLFINNEW
VVLKISTKAIKTLKSKTKLSTFNSNRLHL"
gene
complement (6763..7071)
/ gene="trxm"
complement (6763..7071)
CDS
complement (6763..7071)
/ gene="trxm"
complement (6763..7071)
/ codon_start=1
/ transl_table=11
/product="thioredoxin type m"
/ protein_id="BAC76106.1"
/ db_xref="GI:30409157"
/ translation="MLHIDELTRENVLQSEKLVYDFNAPMGPCRMILIEIAX
ERNLKVQVNTDENPLNLTATFYGIRSLFTLMLFKGQRVDIVIGAVFKSLIHILINKYL
"
gene
7504..8970
/ gene="rbcL"
7504..8970
CDS
/ gene="rbcL"
/ codon_start=1
/ transl_table=11
/product="ribulose biphosphate carboxylase large chain"
/ protein_id="BAC76107.1"
/ db_xref="GI:30409158"
/ translation="MAOSVOERTLKNKRYESCIVIPYAKMWNPNVYVKSIDMLALF
RTPQPGVDPIEAAAVAGESTATWVDTLLTACDVRARAYVDVPAIDQYF
VLYADIDFREESIALNTASTIGNVGFAVAVKALSLDMERPFSLYIKTPOGATCVI
VERRLKRFPRPLGCTVYKRLGSGKNRVRVYBGLKGLDPLKNDENINSPFRKM
RDRVLVMEGVNRAAASGEVKGSLVNTATVECKRAEPRKEVGSVITMLDTIG
YTAIQMAIWARBNMILHIDRAGNSTYSQKNHGINFRISIKMMAGVDIHAGTV

```

Query Match 30.3%; Score 592.8; DB 8; Length 149987;
 Best Local Similarity 60.7%; Pred. No. 9,2e-162;
 Matches 1030; Conservative 0; Mismatches 652; Indels 16; Gaps 3;

```

QY 466 ACCTTACGATGCTTAAAGCACTCCATGATGATACATGATGAGGCGGATATAGCTGCATT 555
DB 80051 ACTTGCCATATGTTAAACATCTTATCTTGTACAGATGCTCGAGATATCTTCGAT 79992
QY 556 GTTACTGAGCTTCCATCTGTTAGACACGGGTGTCGCCGGGCGGTTTGAATGATATT 615
DB 79991 ATTGTTGAGCTTTACATGCTTAACATGCTGCTGCTGCGCCAGCTTAACTGATATT 79932
QY 616 CCCAAGATGTGGCTTAAAGAAATGTAGTACATTCCTCCGACGCCCGGTGACGTTAAT 675
DB 79931 CCTAAGATGTGGATTTAGAAAGAAATGATTAACAGATGCTGTCTCA-----AGTCAAT 79878
QY 676 CTACCGGCTTATCCGCCCAAGGTTAAAGTATCCCGACAAATTAATGGCGCATTCGAA 735
DB 79877 ATCTGATGTCCGCTTATTTTACCACTCTTAACAAATTTAGCAAAATGCTATG 79818
QY 736 TTGTTGAGCAGGCGCAAAATCCCTGCTACGTAAGGAGGAGGCGATCCGCCAAT 795
DB 79817 ATGTTGCAAGAGAGTGTCAACTTTACTTACGTGGGTGAGAGGTGCCAAGC-----T 79764
QY 736 GCGCATGCGCCAGTGCAGAAATTTGGGAAAGTTCCAGTTGCCGCTAACCAACCGCTG 855
DB 79763 GCGTCTCAAGAAATGTATCAATTTAGCTATCTCAATTTCAAAATCTGTAAACCACTAACCTTA 79704
QY 856 ATGGAAATTTGGGCTTTTGAAGAAACCAATCCCTTCCGTTGAGTATGTTGGTATGCAT 915
DB 79703 ATGGAAAGAGTTGTTTGAAGAAACGATCTCTTATTTGGTATGCTTGCATGCAT 79644
QY 916 GGCCACCCCTATGCCAATCTTCCGTCAGCAGAAATGATTTGATTTGCAAGTGGGGCC 975
DB 79643 GGTACGGCTTATGCCAATTTGCTGTAGCAAAATGTATCTTTAATTTGCTTATAGTGCA 79584
QY 976 GCTTTCAGCAGACCGGGTACTGGCAAACTAGAACCAATTTGCTTACCCGCGCAAGTAAAT 1035
DB 79583 CGCTTGTATGATCCGCTACCGGTAACTTATGATGATTTGCTTCTTGCACAGATTAAT 79524
QY 1036 CACATTGACATCGACCCGCGAGGTGGAAAAAAGAGGCTCCGATGTCGATTCCTATG 1095
DB 79523 CATATAGATATGATGCTGCTGAATTCGAAAAATGAAATTCCTCAATTAGCATTCGT 79464
QY 1096 GGGATGATGAGCCCATGTTTGAACAGCTTTGACAGGGGCGCGGAATGATTAACCC 1155
DB 79463 TCTGATGTCAAACTTGTTTAAACAATTTCTTACTGTCACCTTAATTTGCTCTTAATCTT 79404
QY 1156 ACCATCCCATACCAACCCAGACGATGTTAATGCAATTCATTTAGCGGACCGATTAAC 1215
DB 79403 GTTGCTATGATGAACCTATGCTTGGAGAAAGATATGACGTTGAAACAGATTAT 79344
QY 1216 CCCCCTCAGTCCCACTATGAGATATATTGCCCCAGAGGTATGATCAAGAAATT 1275
DB 79343 CCTTACTAATTCCTGCTCTCTTATCTATCCGACAAAGAAATTTAAATGCACTT 79284
QY 1276 GATGCGCAGGCGCCGATGCTTACTACACACCGATGAGGACAAACCAATGTGGGCG 1335
DB 79283 GCGCAATTAGCTCCTCATGCTATATATATGATCAAGTAAACAAATTTAGGCGCA 79224
QY 1336 GCGCATTTTGAACATGCGCCCCCGCGATGATTTTCCAGTGTGCTTGGTATGATG 1395
DB 79223 GCACAGTTATTAAGGTTCACACCCGTCMAAGGCTTTCACGCTTGCTTAAGTACATG 79164
QY 1396 GACTTGTGTTTACCGCGCCATGAGACCAAGTGGAGAGTGGGAGAGCGAGCGTCTC--- 1451
DB 79163 GGCATAGCTTACCTGACGCTATAGTGTGCTCCATCCACATCAAGCAATGATTCAT 79104
QY 1452 ATTTCATCAATGATGATGCAAGCTTCCAAATGATCTTCAAGACTGGGAACCTTAGCC 1511
DB 79103 ATTTCATTTTGGGATGACAAAGTTTCAATATGATTTTACAAAGAACTTGCACATTTCA 79044
QY 1512 CAGTACGATCCAGGTTAAAGTAAATTAATCTCAATAAGTGGGAGGGGATGATGAT 1571
DB 79043 CAATTAAGCTCAAAATCAATCTTCTTTTCAATAGTGTGGCAAGTATGTACGT 78984
QY 1572 CAGTGGCAACAACTTTCTACAGAAAGCTTATTTCTGCTTAAATGATGCCAGGGCATG 1631

```

Db	78933	CAATGGCAACAGCTTCTCTATGGCGAAGAAGATTCACATTTCTTTATGCAAGAAAGTCCA	78922
Qy	1632	CCAGCATTAAATCTCCCTGTGTGAAGCCTATGAGCATCAAGGGTATTACTGTGCCAACCGG	1691
Db	78923	CTGTATTGGAATGCTGTGTAGCAACAAGCTTATGATATCAAAAGCACTCTGTTAGATTCGCA	78864
Qy	1692	GAAGATTGTGGCCCCGGCGATGCGCGAATATGCTAGCCACAGATGTCCTGTGCGATGAGAT	1751
Db	78863	GAACTCTCATGCAACAACATTAGATTCATCTGTGATCTTAACCAAAATGTTTATGTCAT	78804
Qy	1752	GTGTGTGTCAAAAAAGATGAATCTGTACCTTATGATGTCCTCCCGCATGAGTAATGCC	1811
Db	78803	TGTTTAATTCACCCGATGAGAAATTTGATTCCTATGTGTGCTCTCGTAAAGTAATTTCT	78744
Qy	1812	CAATATGCTAGGTTTACCG	1829
Db	78743	CAATATGCTAGGTTCTGTG	78726
RESULT 13			
AF041468			
LOCUS	121524 bp	DNA	1999
DEFINITION	Guillardia theta complete plastid genome.		
ACCESSION	AF041468 X14171 X62349 X51511 X14504 X52158 X52912 X56806 M76547		
VERSION	X62348 Z21976 U81044 AF063017		
KEYWORDS	AF041468.1 GI:3602932		
SOURCE	Guillardia theta		
ORGANISM	Guillardia theta		
REFERENCE	Eukaryota; Cryptophyta; Cryptomonadaceae; Guillardia.		
AUTHORS	1 (bases 47701 to 48415)		
TITLE	Douglas,S.E. and Durnford,D.G.		
JOURNAL	The small subunit of ribulose-1,5-bisphosphate carboxylase is		
MEDLINE	plastid-encoded in the chlorophyll c-containing alga Cryptomonas		
PUBMED	93357429		
REFERENCE	2 (bases 18535 to 19351)		
AUTHORS	Douglas,S.E. and Durnford,D.G.		
TITLE	Sequence analysis of the plastid rDNA spacer region of the		
JOURNAL	chlorophyll c-containing alga Cryptomonas phi		
MEDLINE	DNA Seq. 1 (1), 55-62 (1990)		
PUBMED	92119920		
REFERENCE	3 (bases 43739 to 44938)		
AUTHORS	Douglas,S.E. and Durnford,D.G.		
TITLE	Nucleotide sequence of the genes for ribosomal protein S4 and		
JOURNAL	rRNA(Larg) from the chlorophyll c-containing alga Cryptomonas phi		
MEDLINE	Nucleic Acids Res. 18 (7), 1903 (1990)		
PUBMED	90245597		
REFERENCE	4 (bases 34539 to 35380)		
AUTHORS	Reith,M. and Douglas,S.		
TITLE	Localization of beta-phycocyanin to the thylakoid lumen of		
JOURNAL	Cryptomonas phi does not involve a signal peptide		
MEDLINE	Plant Mol. Biol. 15 (4), 585-592 (1990)		
PUBMED	91338697		
REFERENCE	5 (bases 45872 to 47981)		
AUTHORS	Douglas,S.E., Durnford,D.G. and Morden,C.W.		
TITLE	Nucleotide sequence of the gene for the large subunit of		
JOURNAL	ribulose-1,5-bisphosphate carboxylase/oxygenase from the		
REFERENCE	chlorophyll c-containing Alga Cryptomonas F: evidence supporting		
AUTHORS	J. Phycol. 26, 500-508 (1990)		
TITLE	6 (bases 110917 to 113854)		
JOURNAL	Douglas,S.E.		
REFERENCE	Unusual organization of a ribosomal protein operon in the plastid		
AUTHORS	genome of Cryptomonas phi: evolutionary considerations		
TITLE	Curr. Genet. 19 (4), 289-294 (1991)		
JOURNAL	91330343		

```

PUBMED 1868578
REFERENCE 7 (bases 40675 to 42376)
AUTHORS Douglas,S.E. and Turner,S.
TITLE Molecular evidence for the origin of plastids from a
cyanobacterium-like ancestor
JOURNAL J. Mol. Evol. 33 (3), 267-273 (1991)
MEDLINE 92099311
PUBMED 1757997
REFERENCE 8 (bases 96129 to 98906)
AUTHORS Wang,S.L. and Lin,X.Q.
TITLE The plastid genome of Cryptomonas phi encodes an hsp70-like
protein, a histone-like protein, and an acyl carrier protein
Proc. Natl. Acad. Sci. U.S.A. 88 (23), 10783-10787 (1991)
JOURNAL 92073372
MEDLINE 1961745
REFERENCE 9 (bases 106789 to 108216)
AUTHORS Douglas,S.E.
TITLE A bet homologue is found in the plastid genome of Cryptomonas phi
JOURNAL FEBS Lett. 298 (1), 93-96 (1992)
MEDLINE 92183838
PUBMED 1544427
REFERENCE 10 (bases 42198 to 44153)
AUTHORS Douglas,S.E. and Reith,M.E.
TITLE A bet homolg, encoding a subunit of Mg chelatase, is located on
the plastid genomes of red and cryptomonad algae
J. Mar. Biotechnol. 1, 135-141 (1993)
JOURNAL 11 (bases 82327 to 84799)
REFERENCE 11 (bases 82327 to 84799)
AUTHORS Douglas,S.E. and Murphy,C.A.
TITLE Structural, transcriptional and phylogenetic analyses of the atpB
gene cluster from the plastid of Cryptomonas F (Cryptophyceae)
J. Phycol. 30, 329-340 (1994)
JOURNAL 12 (bases 98901 to 114602)
REFERENCE 12 (bases 98901 to 114602)
AUTHORS Wang,S.L., Lin,X.Q. and Douglas,S.E.
TITLE The large ribosomal protein gene cluster of a cryptomonad plastid:
gene organization, sequence and evolutionary implications
Biochem. Mol. Biol. Int. 41 (5), 1035-1044 (1997)
JOURNAL 97283757
MEDLINE 9137835
REFERENCE 13 (bases 61067 to 68605)
AUTHORS Leitsch,C.E.W., Kowallik,K.V. and Douglas,S.E.
TITLE The atpA gene cluster of a cryptomonad, Guillardia theta: A piece
in the puzzle of chloroplast genome development
J. Phycol. (1998) In press
JOURNAL 14 (bases 1 to 121524)
REFERENCE 14 (bases 1 to 121524)
AUTHORS Douglas,S.E. and Penny,S.L.
TITLE The plastid genome of the cryptophyte alga, Guillardia theta:
complete sequence and conserved synteny groups confirm its common
ancestry with red algae
J. Mol. Evol. 48 (2), 236-244 (1999)
JOURNAL 99128221
MEDLINE 9929392
REFERENCE 15 (bases 1 to 121524)
AUTHORS Douglas,S.E.
TITLE Direct Submission
Submitted (08-JAN-1998) Institute for Marine Biosciences, National
Research Council, 1411 Oxford Street, Halifax, Nova Scotia B3H 3Z1
Canada
COMMENT On or Before Sep 15, 1998 this sequence version replaced gi:11396,
gi:11297, gi:18103, gi:18281, gi:11383, gi:11407, gi:12539,
gi:336730, gi:11300, gi:398949, gi:2661180.
FEATURES
source location/Qualifiers
1..121524
/organism="Guillardia theta"
/organism="Guillardia theta"
/mol_type="genomic DNA"
/db_xref="taxon:55529"
/feature="plastid"
/note="plastid"
/complement(5..86)
/gene="trnY(gua)"
/complement(5..86)
/gene="trnY(gua)"
/product="tRNA-Tyr"
/note="Codon recognized: UAC"

```

gene /anticodon=(pos:50..52,aa:Tyr)
complement(181..573)
/gene="trp119"
CDS complement(181..573)
/gene="trp119"
/codon_start=1
/product="ribosomal protein L19"
/protein_id="AAC35594.1"
/db_xref="GI:3602933"
/translation="MEIILHLYEFVNMKSNKLIQLVESHNSKTLFQLQVDTVRLG
VSTEGNKEIKATGEVITSKNSGNIINTTTRVLQIGIERVYLHSPKLKSIETL
RKSIRAKLYIARTS GKATRLQGRFN"
729..3188
/gene="clpC"
729..3188
/gene="clpC"
/codon_start=1
/product="Clp protease ATP binding subunit"
/protein_id="AAC35595.1"
/db_xref="GI:3602934"
/translation="WPERPTEKAIKYIMLAQSEARRLGHPVGTBOILLIGEGTGI
AAVLKSMGVNLDARVEVKILIGSGFVAVEIPTPPAKRLVLSLEARNQLGHNT
IGTEHLLGIRGEVAAVLENLADLTAVTQVIRLLGDAEVSATNGTKKTP
TLEFSGNLTQKAEGKLPVIGRQKEIRVIOILGRKNPILIGEGVAKTIAE
GLAORINRDPDILEDKRVTLIDGLVAGTKYGEPEERLKKIIDERVANNVILV
IDVHTLIGAGAGAIIDANILKPLALGEMOCTILLEERKHIKDSALERFO
PVYGEPSVRETEIILYGLADRYEKHKVILSPALSAKAPADQYIADRPLDKAID
LIDAGSRVLMNSQLPPARELDKELRLKQKDAVNSQDFETGQRLDRMEIKA
QIAIAHSKRKDEBTEVSVEEDIAQVAATGI PVNKTRESSEKLDQMEETLH
GRIIGDEAVVAASKAIRARVGLKNPNRIASFISGTVGKTELTAALSYFPG
BEAMVRLDMSEYMERHVSRLIGSPGYVYNGQQLTSVRPRPYTVLPBIEKGT
PDVENLLLOILEDGRLLDSKGTVDPRNTLLILTSVNGKVIKEGGGGIFLSEQ
ESQYGRKALVNEIKOYPRPELNRLDEIVPRLTQEVGILAEIMKEVTRISE
KGIQLETTARFKTHLINEGNPIYGARPLRARVRLLEDTLSEFLAEIKIKGTAVV
DVDDGKVKVLSKELVAVN"
3316..3388
/gene="trnf(gaa)"
3316..3388
/gene="trnf(gaa)"
/product="trna-phe"
/note="codon recognized: UUC"
/anticodon=(pos:3349..3351,aa:Phe)
complement(3402..5201)
/gene="dnaB"
complement(3402..5201)
/codon_start=1
/translation="MKVTKEFYNLKAEMHLLIQLIVERDEFLQITRKIDTQVYFQSHQ
ILVRSIHLTKKVLINLPMILSHSELSEIVAVLTLFDRNDPRLININDSLK
VLIDNFIRELONSCAKIVSLTFNPNLSVTLLOKSNLISINTKKVYSLSOL
LLLETILEDKKTNRSTHVLTFGFDLIHVLGKSLPILTAGPSMGKAPMISLVN
LADIQSPPIVIFLSMSSKQILIRLISNENIATSRBENISISEMELINRAMTILS
NLNLTVIDDENLVDLIDQSLASLQIYGGIYGLIVIDYQLDLYKESQSRITELSY
TRYLKLIAQFNLPLVVLSQLSRVFRMLKRPILDLKESGCLSYATQOPFLKSDN
VNFSLTSLKSNHYILSATLELIPQVARIYVIALIRBELQTKYKVVLELDPD
VIVTEHYVLLTNGMKRIELELTVDDLAVLIDQYLIYNNTEVDLFSSNVFESVNL
ICMRIVAVFMT PKTNPFVNALIVNLSI EODADVLFYRDSYCYGNLKNQINDMC
EYIYAKRHGTIGMNLINPNTVSENNLKES"
3594..4073
/note="signal element intein; encodes putative intein
spliced out of the replication helicase subunit"
misc_feature 3594..4073
gene 5361..5433
/gene="trm(cca)"
5361..5433
/gene="trm(cca)"
/product="trna-trp"
/note="codon recognized: UGG"
/anticodon=(pos:5394..5396,aa:Trp)
5588..6013
gene

CDS /gene="trp111"
5588..6013
/gene="trp111"
Query Match 30.3%; Score 592.6; DB 8; Length 121524;
Best Local Similarity 60.2%; Pred. No. 1e-161;
Matches 1018; Conservative 0; Mismatches 669; Indels 4; Gaps 2;
136 ACCGGGGCTTATATCCGATGATGATAGCTGAAAGCCATGGGGTCAAAACATTTTGGC 135
119823 ACAGGTGATTCGCATTAAATGATGATGATGATGATGATGATGATGATGATGATGAT 119822
196 TATCCCGCGCGGCATTTTGGCCCATCTATGATGAATCTGATGATGATGATGATGATGAT 255
119883 TATCCAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 119942
256 GAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 315
119943 TTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 120002
316 GCCAGAGCCACAGGTAAGTGGAGTTGTTGGTACATCTGGACCGGCGGCACTAAC 375
120003 GGAAGATACAGGTAAGTGGAGTTGTTGGTACATCTGGACCGGCGGCACTAAC 120062
376 TTGCTGACCGGCACTGGCCATGATGATGATGATGATGATGATGATGATGATGATGAT 435
120063 TTAGTAACAGGTAAGTGGAGTTGTTGGTACATCTGGACCGGCGGCACTAAC 120122
436 GAGTGGGCGGCGGCACTGATGATGATGATGATGATGATGATGATGATGATGATGAT 495
120123 CAAGTGGGCGGCGGCTTTTATGATGATGATGATGATGATGATGATGATGATGATGAT 120182
496 ACCTTACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 555
120183 ACTTACCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 120242
556 GTTACTGAGCTTTCATCTTCTGATGATGATGATGATGATGATGATGATGATGATGAT 615
120243 GTAGCAGATCATTTTATGATGATGATGATGATGATGATGATGATGATGATGAT 120302
616 CCCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 675
120303 CCMAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 120362
676 CTACCGGCTTATGCGCCCAAGGTTAAAGTATCCCGCAAAATTAATGCGGCTTGCA 735
120363 TTACGAGATGCTCTGCTTAAATAATGATGATGATGATGATGATGATGATGATGAT 120422
736 TTGTTGAGCAGCGGCGAGAAATCCCTGCTACGAGAGGAGGCGGATGCGCGCAAT 795
120423 TTAATAAAGCATCCAGTCACTTTATTAATATGATGATGATGATGATGATGATGATGAT 120482
796 GCCATGCCCAAGTGCAGAAATTTGCGGAAAGTTCCAGTTGCGGTAACCAACCCCTG 855
120483 TCACACACGAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 120542
856 ATGGAATATGGGCTTTGAGAAACCATCCCTTCGAGGATGATGATGATGATGATGATGAT 915
120543 ATGGGTAAGGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 120602
916 GGCACCCCTTATGCCACTTTGCCGTGACGCAATGATGATGATGATGATGATGATGATGAT 975
120603 GGTACTGTTTATGGAATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 120662
976 GGTTCGACGACCGGCTTACGCAACTAAGCAATTTGTTACCGGCGCAAGTAT 1035
120663 CGCTTTCAGACCGGCTTACGCAAGTATGATGATGATGATGATGATGATGATGATGAT 120722
1036 CACATTGACATGACCGGCGGAGTGGGAAAGAAACAGGCGTCCGATGTCGATGTCG 1095
120723 CATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 120782
1096 GGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1155

Db 120783 GG---TGAATCAAAAGACTTTCCTCGTGAATTAATGTAATGTTAAAAATACATTAAAT 120839

QY 1156 ACCCATCCCATACACACCCAGGAGCATGTTAAATCCCATGATCATTTGGCGGACCGATTAC 1215

Db 120840 TTGAGAGGAGACAGAGCCCACTTGGAGTCAACCAATTAAGAATGGCGAAAAAATAT 120899

QY 1216 CCCCCTCCAGTGGCCCACTATAGAACTACTATTTGCCCCCCAGAGAGTATGACACGAAAT 1275

Db 120900 CCACCTCTTTGACAAAATAATTAATTAATTAATCAACAAGAAAGTTATTCACGAAAT 120959

QY 1276 GGTCCGACAGCCCGGATGCTTACTACACACGATGAGGAGCAACACCAATATGAGGCG 1335

Db 120960 AGTACAGAGCTACAAATGCTTATTTTACACAGATGTAAGCCACATCAATATGAGGCA 121019

QY 1336 GCCCAGTTTTTGAACAATGAGCCCGCCGAGATTTCCAGTGTGCTGGTTGGTTACGATG 1395

Db 121020 GCTCAATTTATCAAAAACAAGTCAAAAAGCATGATTAACAAGTGTGTTAGAACATG 121079

QY 1396 GGCTTTGGTTACCTCGCGGCAATGGAGGCAAAAGTGGAGAGTGGGGAGACGAC-GGTCAAT 1454

Db 121080 GCAATCGGTTACCCGACGCTATAGGTGACAAATAGCACATCCAAATGAACAAGTTAT 121139

QY 1455 TGCACTAGTGAAGATGCCAGCTTCCAAATGATCTTCAGAACTGGAAACCTAGCCAG 1514

Db 121140 TGTATCTGTGTGATGCAAGTTCCAAATGAAATATCCAAAGACTAGTACAGTATACAA 121199

QY 1515 TACGACATCCAGGTTAAACTATTTCTCAATTAAGGTTGGCAGGGAGTGGTGGTCAG 1574

Db 121200 TATGATTAACCAATTAATAATTTATTAATTAATAATTAATTAATTAATTAATTA 121259

QY 1575 TGGCAACAACTTTCTACGAAAGCTTATCTGCTTCAATCATGTCCAGGGCATGCA 1634

Db 121260 TGGCAGCAAGCCTTTTACGTAAGAAAGATTTACATTTAATATGAAAAAGTGCTCT 121319

QY 1635 GACATTAATCTCCTCTGTGAAGCTTATGCGATCAAGGGTATTAATCTGTGCGCAAGCGGAA 1694

Db 121320 AATTTTACGAAAGTGGCGAAGCATTTGGCTTACGCTCACTGAATAATAATCCGTTAC 121379

QY 1695 GATTGGCCCGGCGGATGGCGAAATGCTAGCCCAATGCTGTGTGTATGATGATG 1754

Db 121380 GATTCCAACTGAGAAATTAAGAGGCTTTAGATTATGATGACCTATACATGTTGATATA 121439

QY 1755 GTGTCAAAAAAGATGAANAACGTTTACCTTATGATTTGCCCGGATGATTAATGCCAA 1814

Db 121440 CAAGTTATCGCAGATGAANAAGCTTATCCATGTTGCTCCAGGAAAGATTAATGCACA 121499

QY 1815 ATGCTAGGTT 1825

Db 121500 ATGATGGTAT 121510

RESULT 14
AF022186 164921 bp DNA circular PLN 14-DEC-2000
LOCUS AF022186 236235 270297
DEFINITION Cyanidium caldarium strain Rkl chloroplast, complete genome.
ACCESSION AF022186.2 GI:6466296
VERSION
KEYWORDS
SOURCE chloroplast Cyanidium caldarium
ORGANISM Cyanidium caldarium
Eukaryota; Rhodophyta; Bangiophyceae; Porphyridiales;
Porphyridiaceae; Cyanidium.
REFERENCE
AUTHORS Vogel,H., Fischer,S. and Valentin,K.
TITLE A model for the evolution of the plastid sec apparatus inferred from secY gene phylogeny
JOURNAL Plant Mol. Biol. 32 (4), 685-692 (1996)
MEDLINE 97134960
PUBMED 8980520
REFERENCE
AUTHORS Glockner,G., Rosenthal,A. and Valentin,K.
TITLE The structure and gene repertoire of an ancient red algal plastid

JOURNAL genome
MEDLINE J. Mol. Evol. 51 (4), 382-390 (2000)
PUBMED 20496959
PUBMED 11040290
REFERENCE
AUTHORS 3 (bases 46857 to 47851)
TITLE Valentin,K.
JOURNAL Direct Submission
SUBMITTED (22-MAR-1996) Institute for Plant Physiology, Justus Liebig University, Heinrich Buff Ring 58-62, Giessen 35392, Germany
AUTHORS 4 (bases 28701 to 75580)
TITLE Gloeckner,G., Rosenthal,A. and Valentin,K.
JOURNAL Direct Submission
SUBMITTED (02-SEP-1997) Department of Genome Analysis, IMB Jena, Beutenbergstr.11, Jena 07745, Germany
AUTHORS 5 (bases 1 to 164921)
TITLE Gloeckner,G., Rosenthal,A. and Valentin,K.
JOURNAL Direct Submission
SUBMITTED (18-NOV-1999) Genome Analysis, Institute for Molecular Biotechnology, Beutenbergstrasse 11, Jena 07745, Germany
REFERENCE 6 (bases 130696 to 132364)
AUTHORS Vogel,H., Fischer,S. and Valentin,K.
TITLE Submitted (18-NOV-1999) Institute for Plant Physiology, Justus Liebig University, Heinrich Buff Ring 58-62, Giessen 35392, Germany
COMMENT On or before Nov 23, 1999 this sequence version replaced gi:529651, gi:1240002, gi:2465730.

FEATURES
Location/Qualifiers
1..164921
/organism="Cyanidium caldarium"
/organelle="plastid:chloroplast"
/mol_type="genomic DNA"
/strain="Rkl"
/db_xref="taxon:2771"
/join(164808..164921,1..78,462..644)
/note="direct repeat"
/repeat_unit separated by stem loop"
339..477
/note="stem-loop separates direct repeat unit"
580..780
/note="similar to yeast mitochondrial origin of replication region in Genbank Accession Number U36902"
986..2041
/gene="psbD"
986..2041
/gene="psbD"
/product="unknown"
986..2041
/gene="psbD"
/note="Photosystem II D2 protein"
/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="AAF13026.1"
/db_xref="GI:6466445"
/translation="MTAIGREQERGWEDLDMDLKRDFVFIQWSGILLPPCAVIAL
GAMFTGTPVSWYTHGLASYLEGCEPFLTAIVGSPANMGSLFLIMPEACQEPTR
WCOIGLITFTALHSGFGLICLROFETARLVGAPYVAIAFSGPIAVFVPLLYP
LQASWEPAPSFQVAIAIPRFLIFLOGFHWNTLNPMMGVAGIILGALLCAHGTAE
NTLPEGASDTPRAFTPROSEETYSMTANRPMQIQRVAIRANKRMILHFLIPVPT
GLAVSSIGIVGALNIRAYDFVQSIRAEDEFEFTYKILLNKGIRAWAADQP
HENFVPEEVLPRGNAL"
2031..3413
/gene="psbC"
2031..3413
/gene="psbC"
/product="unknown"
2031..3413
/gene="psbC"
/note="Photosystem II 44 Kd apoprotein"
/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="AAF13025.1"

```

/db_xref="GI:6466444"
/translation="MLFNKDKNNGIGRSISTGFAMWSGNARLINLSGKLLGHAHVA
GLIVFTGAMTLFETSHPEKPLYEQGMILPHLATLGWVAGGEIVNTYPTFGT
VHHVSAVYFGGIVHSIVGPDVLTSPFSPRYDMDRDKKMTTIGIHLILIGGAF
LIVTKALFTGCIYDITAPGGADIRFTINPLNPAIIIPSYLXSPFGEGEMIVGNAMF
DVIIGHWIVTCVIGIWHILTRPESWARRAVWSGEALISYLGALMGGTAAEY
AWNNIVYSEFPGPTAAEASQAQAFPIVDRQAGANLASTGPTGLKYLMSRSTG
EVLIGETMRFDLRAFWLEPLRSSNGLDNKIKNDIQPQERRAAEVYTHAPLGSLN
SVGGATEINSVNVSPRSWLTTSHPFGPIFGIHLMHGARRAANAAGFEKGINEN
EPLSMRPLD"
3546..3821
/gene="ycf19"
/note="similar to ycf19 in Porphyra purpurea"
3546..3821
/gene="ycf19"
/product="unknown"
3546..3821
/gene="ycf19"
/codon_start=1
/transl_table=1
/product="unknown"
/protein_id="AAFL3024.1"
/db_xref="GI:6466443"
/translation="MKDGTDLIETTAPQIYIYIILMLSLGMPENINWYSQPFYS
LSQSDPYNLPHGVPPSPFGIDSPPIGTLIDFIIEILSRQLXDL"
3873..4133
/gene="ccr19"
3873..4133
/gene="ccr19"
/product="unknown"
3873..4133
/gene="ccr19"
/codon_start=1
/transl_table=1
/product="unknown"
/protein_id="AAFL3023.1"
/db_xref="GI:6466442"
/translation="MNNCLDPKNKVFEXILNFKPCIDPWCNTDLFKFPLVYKLNKKKQFT
NKFYKMKKKYIDIMLIFYIYHIVIRIGKQTKDLQITKC"
4189..4272
/gene="ccr10"
4189..4272
/gene="ccr10"
/product="unknown"
4189..4272
/gene="ccr10"
/codon_start=1
/transl_table=1
/product="unknown"
/protein_id="AAFL3022.1"
/db_xref="GI:6466441"
/translation="WTKLKSLLNDKULTLISLUNKIYIDIN"
4334..4406
/product="tRNA-Thr"
4450..4971
/gene="ycf3"
/note="similar to ycf3 in Porphyra purpurea"
4450..4971
/gene="ycf3"
/product="unknown"
4450..4971
/gene="ycf3"
/codon_start=1
/transl_table=1
/product="unknown"
/protein_id="AAFL3021.1"
/db_xref="GI:6466440"
/translation="WPRSQRKDNFIDTFTSIADLIYKILPTKESKEAEVYVYKDKKA
AOSGEVABALACYQAKTKEDPMDSFTIYVNGIQAASNGARLERYHESLKEN
ENLVQALNNIVAVIHYGNKLFQSKLQEAALMPDKASNNWRKAIKLAIPNYIEADNW
LKITGRITIEDIML"
5004..6476
/gene="atpb"

```

```

mRNA
5004..6476
/gene="atpb"
/product="unknown"
CDS
5004..6476
/gene="atpb"
/note="ATP synthase CF1 beta chain"
/transl_table=1
/codon_start=1
/product="unknown"
/protein_id="AAFL3020.1"
/translation="MKPQOLKQKQTLAIKVNKKGVIOIGPVLIDIVPDGOLPKV
FNATKINNSNNMTTCCEOILGDKXRAVAMSTEGLRGASALIDGEPISIPVGE
TLGRIFNLGPIPEKEGPVINDKLPPIHRAPKETQCFEPKDSIFENGITKVIDLAPVR
RGKIGLPGAGVAKTVLMELINNVAKHAGSVFPGVGERTRBGNDIYQEKESGV
INERDLNLSKVALCYQMNEPFGARMEVGLTVAEYFRDVAKQNVLLFTDNTFRV
QAGSEVSLALGRMPSAVQYQPTLIGTEGALQERITTSIDISITSIQAVVPADLTD
APATFPAHLDATVLSRALAKGIYPAVADPESITMLQGISDEHYTTARKVETL
QRYKELIODIITAILGDELSEEDRLIVSRARXIEFISLOPFVAEVFGISGKYVSLSD
SIKGFMTLSGEVNIIPQAFVLYIGRIEBAIDKAKQVEKS"
6486..6887
/gene="atpb"
6486..6887
/gene="atpb"
/product="unknown"
6486..6887
/gene="atpb"
/codon_start=1
/transl_table=1
/product="unknown"
/protein_id="AAFL3019.1"
/db_xref="GI:6466438"
/translation="MALTVKVTDPDVRVWVKTVVEIILPSGTQGLIMNHAFLTAL
DIGVRAMNTWVPLVILGFAOIDNNLVITIIYSDBEKVKAIDEBEANKLLAASLNL
Query Match      29.8%; Score 584; DB 8; Length 164921;
Best Local Similarity 59.7%; Pred. No. 3,56-159;
Matches 1019; Conservative 0; Mismatches 680; Indels 7; Gaps 2;
134  AAACGGGGGCTTATTCCTGATGATGATACCGCGAAGCGCATGGGGTCAACACATTTTGG 193
Db    73843  AACCTGGGGCCCTTTCCTTAATAGATGCTGCGTAAACAAAGTAAATATTTTGG 73902
Qy    194  GCTATCCGGCGGGCGCAATTTGGCCCATCTATGATGAACCTGACCGCTTTGAACGGCGG 253
Db    73903  GATATCCAGAGGAGACGATTTCTCCCATCTATGATGAACCTATACATCGGAAAAAAGA 73962
Qy    254  GGGAAATTGAGCATTTTTCGTGCGCCATGAACAAGAGCTTCCCATGCGGCGGATGGGT 313
Db    73963  AACTAATTAAACACTATTTAAGTTCGTCAAGAAAGGCGAGCTCATGACAGCATGCTT 74022
Qy    314  ATGCCAGAGCCACAGGTAAAGTGAGAGTTTGTTCGTACATCTGACCAAGCGCGACTA 373
Db    74023  ATGCCAGAGCACTAATGAAGTAGGGGTGTGTTAGCTACATCTGGGCGTGGCCGACAA 74082
Qy    374  ACTTGGTAGCGGCAATTCGCAATTCGCAATTCGCAATTCGCAATTCGCAATTCGCA 433
Db    74083  ATTTAGTAACAGAAATGACTGCTCAAGAAAGGACTGTGTTCCATTAATTCCTATTAG 74142
Qy    434  GAGAGTGGGCGCGTCCAGATGTTGGTAGAGTCTTCCAGAAATGACATTTTGGCA 493
Db    74143  GACAAAGTAGCAGACATTTATAGTAGCATGCTTTTCAAGAAATGATTTTGGAA 74202
Qy    494  TCACCTTACCGATCGTTAAGCACTGCTATGTGTAACGTAAGTGGCGGATATGCTCGCA 553
Db    74203  TTACACTACCGATTTAAACATTCATTCGTAGTAAGAGACCCAGATATATTCACAA 74262
Qy    554  TTGTTACTAGAGCTTTCATCTGTGTACACACCGGCGCGCGCGGCTTTTATGATGATA 613
Db    74263  TCGTTTCAGAACCATTCATATTTAGTAACCATGTAAGCACAGGAGCTGTTTATATGATG 74322
Qy    614  TTCCCAAGGATGTGGGCTTAGAAGATGAGTACATTCCTCCGACACCC-----GGTG 667

```

Db 74323 TACCTTAAGATGTTGTAAGTAAGTAATTAATCATATGCTAATCCGTGATC 74382
Qy 668 ACCTTAATATACCGGGTATACGCCACCGTTAAAGTAATCCCGACAATTAATGCGG 727
Db 74383 ACAACCTAATTAATACAGGCCAATATAGCGGCTTGATGATTAACAATTTGAAAAAT 74442
Qy 728 CATGCAATTTGTGAGCAGCGCAATATCCCTTGTCTACGTAGGCGGAGCGGCTGATCG 787
Db 74443 TCAAAAAATGCTTTTAGAGTCMAACAGCTAATCTATATGTGAGCGCGGCTGTATA 74502
Qy 788 CCGCCATGCGCCATGCGCCAGTGCAGAAATTTGCGGAAGGTTCAGTTGCGGCTGATACA 847
Db 74503 TGTCTAGAGCTACGACGAAATGAGAACTAGCAAGCTTTATMAAAATTCGGGTACCA 74562
Qy 848 CCAACCTGATGGGAATTTGCGGCTTTTACGAAAAACCATCCCTTCGGTGGGTATGTTGG 907
Db 74563 CAACCTTAATGGGGAAGGATCTTTTAATGAATACATCTCTTGTATCTTGGAAATGCTTG 74622
Qy 908 GTATGATGCGCAACCGCTATGCGCAACTTTGCGGTACGGAATGTGATTTGTGATTCAG 967
Db 74623 GCATGATGGAATCTGTTATGGAATCTTGCTGTATGCGAGTGCATTTGCTCATTTGCTT 74682
Qy 968 TGGGGGCGCGTTTCGACGACCGGCTAACTGGCAAACTAGACGAATTTGCTAGCCGCGCA 1027
Db 74683 TAGGGGCAAGATTTGACGACAGAGTGAATGGAATTTGATTTGATGCAATGCTC 74742
Qy 1028 AAGTAATTTACATGATGATCGACCGGCGGAGTGGGAAAAACAGGGCTCCCGATGTCG 1087
Db 74743 AAGTAATTTACGTCGATGATGATGCTGCAAAATTTGGCAAAATTCGCAATTG 74802
Qy 1088 CCATTGTGGGGATGATGACCCATGTTTGAACAGCTTTTTCAGCGGGCGCGGAATGG 1147
Db 74803 CATTAATTTACATATTAATAATAGTGTGAAAGATTTATGCTGTATGAAAGAGAA 74862
Qy 1148 ATTATCCCAACCATCCCATACACCGGCAATGTTAATTCGATTCATTTGCGCGA 1207
Db 74863 CAACCAATTTGATTAATAATCAACACAAGCTTGCTACATGATTCATAATGAGA 74922
Qy 1208 CCGATTAACCCCTCCAGGTGCGCCCATGAGAGATCATTTGCCCGCCAGAGTATGAC 1267
Db 74923 AAGAATATTCATTAAGCATCTCTCATGATGATTAATCTTATTCACAGAGATTA 74982
Qy 1268 ACGAAATTTGTGCGCAGCGCCCGCATGCTTACACACGCGATGTGGAACAACCA 1327
Db 74983 ATGAATATGATCAATATAGCAAAAGGCTTTCTTTGCAACAGATGTCGCTCAACAT 75042
Qy 1328 TGTGGGCGGCGCGAGTTTGAACAATGGCCCCCGCATGATTTCCAGTGTGCTGG 1387
Db 75043 TGTGGGCTGCTCAATTTTAAAGTTGAACAAGCAAAATGCTATCAAGCTCTGCGCTTG 75102
Qy 1388 GTACGATGGGCTTTGCTTACCTGCGCCATGAGGACCAAAAGTGGAGTGGGGGACGAGC 1447
Db 75103 GTACGATGGGATATGATGATGCTGCGCGGATGAGGCAAAATTTCTAATCTTAACGACC 75162
Qy 1448 -GCTATTTGATCGATGAGATGCGACGCTTCAATATAATCTTCAGGAATCGGAAACC 1506
Db 75163 TAATTAATATGATTAACAGAGACGCGAGTTTCAATGAACTGTGCAAGAACTAGGGACTA 75222
Qy 1507 TAGCCCAAGTACGATCAAGGTTAAATCTATATCTCAATTAAGCGTTGGCGAGGAGATG 1566
Db 75223 TAGCAATATATAGTATGATATCAAAATTTATTAATTAATTAATAGCAAGGATGG 75282
Qy 1567 TGGCTGATGGCAAAAATTTCTACGAAGAACTTTATCTGCTTCTAATATGTCGCAAG 1626
Db 75283 TCAGGCAATGGCAACAGCTTTTTCGACCGAGGATGCGCACCTTAACATGCGCAAG 75342
Qy 1627 GCATGCGAGACATTAATCTCTCTGTAGAGCTATGAGATTAAGGATTAATCTGTGCCA 1686
Db 75343 GACAACTGATTTTGTATTAATAGCAAACTCTTACGGGATTAAGAGATTAAGATCAAGA 75402
Qy 1687 AGCGGAAAGATTTGGCGCGCGATGCGCGCAATGCTAGCCCAATGCTCTGTGTGA 1746

Db 75403 CAAGCAAGATTTAAATCAAAAATTTGAAGAAATTAATATGACACACTGCGCCCTTTGTGA 75462
Qy 1747 TGAATGTGGTGTCAAAAAGATGAAAAAGTTCACCTATGATGATGCCCCCGGAGATGCTA 1806
Db 75463 TCGATGATGCTGTGCTGCTACGCAAAATTTGTTATCCAAATGATGCTCTGGAAGAGTA 75522
Qy 1807 ATGCCCAATGCTAGTTTACCGGAA 1832
Db 75523 ATTCTCAATGTTAGACTCAAAA 75548

RESULT 15
LOCUS AE017162
DEFINITION Prochlorococcus marinus subsp. marinus str. CCMP1375 section 2 of 6 of the complete genome.
ACCESSION AE017162
VERSION AE017162.1 GI:33237292
KEYWORDS
SOURCE Prochlorococcus marinus subsp. marinus str. CCMP1375
ORGANISM Prochlorococcus marinus subsp. marinus str. CCMP1375
Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcaceae; Prochlorococcus.
REFERENCE 1 (bases 1 to 300355)
AUTHORS Dufresne, A., Salanoubat, M., Paterinsky, F., Artiguenave, F., Kham, I. M., Barbe, V., Duprat, S., Galperin, M. Y., Koonin, E. V., Le Gall, J. F., Makarova, K. S., Ostrowski, M., Oztas, S., Robert, C., Rogozin, I. B., Scanlan, D. J., Tandeau de Marsac, N., Weisenbach, J., Wincker, P., Wolf, Y. I. and Hess, W. R.
Winker sequence of the cyanobacterium Prochlorococcus marinus SS120, a near minimal oxyphototrophic genome
Proc. Natl. Acad. Sci. U.S.A. 100 (17), 10020-10025 (2003)
2 (bases 1 to 300355)
REFERENCE
AUTHORS Dufresne, A., Salanoubat, M., Paterinsky, F., Artiguenave, F., Aumann, I. M., Barbe, V., Duprat, S., Galperin, M. Y., Koonin, E. V., Le Gall, J. F., Makarova, K. S., Ostrowski, M., Oztas, S., Robert, C., Rogozin, I. B., Scanlan, D. J., Tandeau de Marsac, N., Weisenbach, J., Wincker, P., Wolf, Y. I. and Hess, W. R.
Direct Submission
Submitted (28-May-2003) UMR 7127 CNRS et Universite Pierre et Marie Curie, Station Biologique de Roscoff, Place Georges Tessier BP74, Roscoff 29682, France
FEATURES
source location:Qualifiers
1..300355
/organism="Prochlorococcus marinus subsp. marinus str. CCMP1375"
/mol_type="genomic DNA"
/strain="CCMP1375; SS120"
/sub_species="marinus"
/db_xref="taxon:167539"
86..706
/gene="trxA"
/locus_tag="Pro0315"
86..706
/gene="trxA"
/locus_tag="Pro0315"
/codon_start=1
/transl_table=11
/product="thioredoxin family protein"
/protein_id="AA099361.1"
/db_xref="GI:33237293"
/translation="NWKTSITMLPLGYAMPSPDLPILQINDTLTFSSSNTAARFTSM LNKPLIMICAHCPFVNIESLSKLDQDFDIOPLAVASNSVSHPEDGPEFLV QOIIRKGNAPYLLDLEOTPAKSLKACTPFFLFSYPMKGSNHLIYRQQLDSRRCN NIDPNGFDRRLNVLNATVNSCEVPIDQKPSIGCNIKRBEFSFWSWG"
769..1956
/gene="wecF"
/locus_tag="Pro0316"
769..1956
/gene="wecE"
/locus_tag="Pro0316"
/codon_start=1
/transl_table=11

/product="DegT/DnrJ/EryC1/Stx5 aminotransferase family
enzyme"
/protein_id="AAP9362.1"
/db_xref="GI:33237294"
/translation="MQYPPPTLEKQSLSTIGQDDDEAVSRVYKSGIYIGSGEYLKFEROS
FADAVGSPFALIGNSGTDALVALRALNIGPMDEVTTPSPFATKABASNGATPEF
VVDVPTNYLDELIELAITEPTKAILPVALFGCPVMTKIIVSLAKNGIKVIEDAQ
AVGTQMDGKGVSGIGDIGCFSPFTKVLGAAGAGATTNDPELARIJELAIHGPFC
RYFTEIGNSRLDAVOAILNVALPLIDKMTIQKREIARVYNLKGISDLAPFC
EKNNDHSMQFVIRVADKSGEEKVKGIKKIENIKDLPESDARDLTKRLLHDGVNT
IIVYPIPIHQPAKLGKVKNNLPNTSLCSQVLSIPFELTLEGEVYVNSIKS
LKD"
/complement(1994..2554)
/locus_tag="Pro0317"
/complement(1994..2554)
/locus_tag="Pro0317"
/codon_start=1
/transl_table=1
/product="NUDIX/MutT family pyrophosphatase"
/protein_id="AAP9363.1"
/db_xref="GI:33237295"
/translation="MSPLSQOEPEKILENIQALDSKKIREINRFEPLIGVECTFGVI
KHGPAALAVPIKENGAVIVLROYFAISRIIEFGPAGTLPNEAPLKTIORELOESG
YKAKMDSLGVWMPGPGVDEIILHYLARETEPLIKOPSGDEDEIEVLMTSELED
CIASGNALDSKSVTAMHACOVLPK"
/complement(2626..3207)
/locus_tag="Pro0318"
/complement(2626..3207)
/locus_tag="Pro0318"
/gene="folK"
/codon_start=1
/transl_table=1
/product="7,8-dihydro-6-hydroxymethylpterin-pyrophosphokinase"
/protein_id="AAP9364.1"
/db_xref="GI:33237296"
/translation="MSSINYLISKTLAIGGANIESPAGSPAKTLITAPRLLENTTS
WIKSIXKASXNSTEDSTIDPSMELVTRTAIDGPKTOPFINALVIRRGINOLIP
SVKARDLNLLILIEKFRGRDRISKIHNGPRSLDILAMGDLQINBSDLILPHPY
LIERDPVILPLAQALVKSSSPIQIPQKGMPE"
/complement(3255..5378)
/locus_tag="Pro0319"
/complement(3255..5378)
/locus_tag="Pro0319"
/gene="chld"
/locus_tag="Pro0319"
/EC_number="4.99.1.1"
/codon_start=1
/transl_table=1
/product="Protoporphyrin IX Mg-chelatase subunit Chld"
/protein_id="AAP9365.1"
/db_xref="GI:33237297"
/translation="MWVPSKEKEPOSISISVLPSELEANTSPFLAAVTHGHTLKXLMMA
AVDPGLGVIIAGRGTEKSYLARGHLHLLPILVLDLRFGDLOGRFPRGRNLDPRF
PEEMDANTELIKORESNNSANGKODLOTRVOAPFVQVPLGTEDELVAIVASSL
DSGSIVFQPGFLAAHARGVLIIDELNLDGIIVMLAAVSGENQVERGLSLSHC
RPLIATVPEEGAVRDLDRFAIVLSADQIISNEQREIETESLAFSGSTKPSK
MDEATQALATOLLAROMLSDIALSOGIETVLKRAIRGVGHRSELVARYAKAKA
ALCGRDKYAEADILVAVRIIVAPRSLOIPPEROMPEPDEOQPPRESIDDEDE
DEDEDDMPEDQAPPIPEEFMLDPKCAIDPULLPSAKSKSSGSGSRSAVFSNS
RGRVYKPTLPKGPVTRIAVDAITLRAAPYQAREREPGRVILVQEDLAKLQKKA
GSLVFLVADGSMALNMQAKGALRLTEAVESRDEVSLIPRGEQVALLPPTIR
SITAKRLLETPCGGSLAHLGLOARVGNALSTGDLQOVVVAITGGRNIPIG
KSLGPELIDGDSVLDKOEVLDIATRYRTGIIKLIVDTERKFIASGIGDLADAAG
KYVOLPKATDOLASIMDAINTVK"
/complement(5387..6232)
/gene="ltg2c"
/locus_tag="Pro0320"
/complement(5387..6232)
/gene="ltg2c"
/locus_tag="Pro0320"
/codon_start=1

/transl_table=1
/product="ABC-type transport system involved in resistance
to organic solvents periplasmic component"
/protein_id="AAP9366.1"
/db_xref="GI:33237298"
/translation="MRRSFRPAIYGFSLIGSVIFSGLTMLKGLKXISNMTWYFANF
ADAGSLGCTVTRGQIGGVKIIITSSQVPAKIRINNUNVLPFDPVYAKIBETSL
LGGDVEENIQSGPILSISNPKQZCPDLICADSDKGNLDNLSLTGGINPF
LNEAENEBIKRINSLEQPDSTENDELIRSLKELFPARPIIDELKTVAHINI
LGSIDDPILSDISSSSISKSPFAKDKISGVKDEIINDELITNGIKKAIIGIKL
NDIYE"
/complement(6238..7050)
/gene="ltg2a"
/locus_tag="Pro0321"
/complement(6238..7050)
/gene="ltg2a"
/locus_tag="Pro0321"
/codon_start=1
/transl_table=1
/product="ABC-type transport system involved in resistance
to organic solvents ATPase component"
/protein_id="AAP9367.1"
/db_xref="GI:33237299"
/translation="MKDLPSKPISSKVPYVENDSLMKWADNSLVNKVSIKMPGER
IAILGPSGCGSTVRLRLAGLLPSGDLKFLGMONYLRIDQTFPPDVRILVFQNPAL
IGSLVEENVGFLRKNSITDKRAREIVISCLEVEGLYNIDKFPQLSGGMQRVS
FARALINDPKSDSMPLELFDDEPTAGDPVACTRIEDLYKTTLLAMGCSIVVSHW
STIERSARVMVLVGEFQWMDGSIIDPFKNTQNPPIIKQFRITSGEPQKXI"
/complement(7227..8555)
/locus_tag="Pro0322"
/complement(7227..8555)
/locus_tag="Pro0322"
/codon_start=1
/transl_table=1
/product="Uncharacterized conserved membrane-associated
protein"
/protein_id="AAP9368.1"
/db_xref="GI:33237300"
/translation="WLGRIOQALGMLPGLVYKEMTSLGILVYALFGASIMWDLRP
IYVVEGIFLIGITIKFLPSITGPVIFVIGISLLIWGQSRFSKSIONALAPKDDV
LIDARAKSKNKGNTVYAGGTGSLALGLKARYSRIYATVADGSGSGILRL
ELGVPPGDINDCLAAISTERSLITRLPQVPSGCTGVGHSFGNLTALTSITGNL
ETAITASSRVLAVOGVVPAITNADVRLMAELNENBRIGESALINVSPIRIGCYPE
QPPALPSAIEAIEENELVIDPGSLYLSLLPNLILPELVEALIQSKAKKITCNLMTQ
FGETDGLDIAGHRAIEAQLASFGVTRIFNSIIAQKELPSALLEYLSGABPVIC
DNRRLRSGGYRLFKASLOESKTISSKSVLRHDPKILSLAVMPPYKRYKDA"
/complement(8556..9101)
/gene="ndhJ"
/locus_tag="Pro0323"
/complement(8556..9101)
/gene="ndhJ"
/locus_tag="Pro0323"
/codon_start=1
/transl_table=1
/product="NAD(P)H-quinone oxidoreductase subunit J"
/protein_id="AAP9369.1"
/db_xref="GI:33237301"
/translation="MNEBLDKKITTEETNQVLEPORGFISLLCKSGIEHTVNEYDHI
GIEIISVPEFLDIIIVSLKNGFNVLCCQGYBGPGLNTVCFYHLIATINYSKDSK
PEBVRVYFLNNDGDLKVPISYKIFRSGDMQGERETVDMFGVAFDHPKRLMBEDW
KGMPLRKXIOPDPYEMODAY"
/complement(9104..9832)
/gene="ndhK"
/locus_tag="Pro0324"
/complement(9104..9832)
/gene="ndhK"
/locus_tag="Pro0324"

Query Match 28.9%; Score 566; DB 1; Length 300355;
Best Local Similarity 59.3%; Pred. No. 7.5e-154;
Matches 1004; Conservative 0; Mismatches 680; Indels 10; Gaps 2;
QY 139 GGGGCTTATATCTGATGATAGCTGAACGCGCATGGGGTCAACACATTTTGGCTAT 198

```

Db 195975 GGGGCGAGATGCGCTTATGGAAATCTCTTAAATAFCAGGTGTCGATACATATTTTGGGTAC 196034
QY 199 CCGGCGGGGCAATTTTGGCCATCTATGATGTAAGTACCGCTTTGAAGCGGGGGAA 258
Db 196035 CCGGAGGAGCTATCTCTCTATCTATGATGCGGTTTAAAGCAAGAAAAGCAGGTGG 196094
QY 259 ATTGAGCATATTTTGGTGGCCATGAACAAGAGCTTCCCATGCGCGCGATGGTATGCC 318
Db 196095 GTAAACATATTTTATGATAGACATAGCAAGGTGGTGTCTCATGCGCTGATGATGCA 196154
QY 319 AGAGCACAGGTAAAGTGGAGATTGTTGGGTATCATGTGACCGAGGCGGCGACTACTG 378
Db 196155 AGGTCCACAGGAAAGGTGGTGTATGTTTGGCACCTTAGAGCCAGAGCAACAAATTTG 196214
QY 379 GTGACCGGCAATGCGCAATGCCATTTGGACTCGGTGCCATGGTGTATTAAGAGAG 438
Db 196215 GTTACAGGAATGCCACAGCTCAAAATGGAATCTGTTCATTTGTTTATTACTGACAA 196274
QY 439 GTGGGCGGTGCCATGATTTGTAGAGATGCTTCCAGAAATTTGACATTTTGGCATGCC 498
Db 196275 GTCCCAAGACAGCAATTTGGAATGACGCTTTCAAGAAACAGATATTTTGTATTAAT 196334
QY 499 TTACCGATCTTTAGCACTCTCTATGTGTAGCTAGTGGCGGAGATGGCTCGCATTTGT 558
Db 196335 CTCCCAATTTGTAACATTTCTTGGGTGTAGAGATCTTCTGAAATAGCATCTGTATTT 196394
QY 559 ACTGAGGCTTTTCATCTTGTCTAGCACCGGTGTGCCGCGCGGTTTGTATGATTTGCC 618
Db 196395 GCGCAGGCTTTTATTTATGCTTCTCTGCGGCCAGAGACCAAGTCTGTATGACATTTCA 196454
QY 619 AAGATGTGGGCTTGAAGAAATGTAGTACATTTCCCTCGACCCCGGAGAGTTAATCTA 678
Db 196455 AAGATGTGGGCGCAAGAACTTTTATATATCACTGTTCAGCCAGGTTCTTTATCTCT 196514
QY 679 CCGGCTTATGCGCCCAAGGTTAAAGTAAATCCCGCAAAATTAATGCGGATTCGAATGG 738
Db 196515 GAGGCTTATCTCTCTCTCAAAAGCCAGACGAGTATATCTAAAGCCTGATTTA 196574
QY 739 TTGAGAGAGGCGAGAAATCCCTGTGCTACGATAGGAGGAGGGCGATGGCCGCAATGCC 798
Db 196575 ATTGAAGAGCTTAAAGGCACTTTTATATATGAGAGAGAGATTAATTTGCAAGTCT 196634
QY 799 CATGCCAGGTGACAGAAATTTGCGAAGGTTCCAGTTGCCGATACCAACACCTGATG 858
Db 196635 CATGACACTTATTTGCTATTTGCCAAGGATATGACTTGGCTGTATGACAACTGATG 196694
QY 859 GGAATTTGGGGCTTTGACGAAACCATCCCTTTGCGGTGATGTTGGGTATGCAATGCC 918
Db 196695 GGTAAAGGTGCAATTTGATGAACGTAATCTTTCTGTGGGAGTTTAAAGGATGCATGCC 196754
QY 919 CACCGCTATGCCAATTTGCCGTACGCAATGTGATTTGTGATTTGCAATGGGGGCCGT 978
Db 196755 ACTGCATATGCCAATTTGCGAGTAACTGAATGCAATTTATAGTATAGGAGCAGCT 196814
QY 979 TTGACGAGCGGCTTACTGGCAACTAGACGAATTTGCTAGCCGCGCCAAAGTATTCAC 1038
Db 196815 TTGACGAGCGAGTACTGGAAGACTTGTATTCATTTGACGCAAAATGCAAAAGTTGTCAT 196874
QY 1039 ATTGACATGACCCCGGAGGTGGGAAAAACAGGGCTCCGATGTGCGCATTTGTGGGG 1098
Db 196875 TTGGAAGTGAACCTGCGAGAAATTAACAAAACAGAGTTGCAAGTGTGGCGTTTAAAGA 196934
QY 1099 GATGTAGCGCATGTTTAAACAGCTTTTGACGCGGCGCCGGAATTGATTAACCCACCC 1158
Db 196935 GATTTAGCAATTAAGTCTTGTGATTTACATGAACTTAGCAAAAGAGA-----ACT 196985
QY 1159 CATCCCATACCAACCCAGGCAATGGTTAAATGCAATTTGATCATTTGGCGGACCGATTAACCC 1218
Db 196986 GTTAAACCAATATCTCAAGAAATGGTTGGCTCAAAATAGAGAGTGAAGATTAATTAATCA 197045
QY 1219 CTCGAGGTGCCCACTATAGGATTAATGCCCCCAGAGGTAGTACAGAAATTTGGT 1278

```

```

Db 197046 CTTTAAATACCTTCAAAAGAGGGAAATTAATACCAACAGAGTATGATTAATTAAG 197105
QY 1279 CGCAGGCGCCCGCATGCTCTACTACCAACCATGTGTGGAGCAACACCAATATGTGGCGGCC 1338
Db 197106 GATTTAGCAACAGATGCTTATTAACACAGTGAATGGGCCAATCAATATGTGGCAGCT 197165
QY 1339 CAGTTTGAACAATGGCCCCCGCCAGATGATTTCAAGTGTGGCTTGGGTGAGATGGCC 1398
Db 197166 CAATATTTAAGAAATGTCCACGCTCATTTGATTAAGAGTGCAGAGATTTGGGAAACAAGGGT 197225
QY 1399 TTTGGTTTACCTGCGGCATAGGAGGCAAAAGTGGAGTG- GGGAGAGAGCGGTCAATTTGC 1457
Db 197226 TTGGAATGCTCTGCTCTTAGAGTGCAGATTTGCTTTGCAAAACAAACAGATTTTGC 197285
QY 1458 ATCAGTGAAGATGCCAGCTTCCAAATGAATCTTCAAGAACTGGAACTCTAGCCAGTAC 1517
Db 197286 ATTGACAGTGAATGCAAGCATTTTATGATGAATTAATCAAGAAATTAAGGACTATATCTGCAATAT 197345
QY 1518 GACATCCAGTTAAATCTATTAATTCATTAACGTTGGCAGGGAGATGTGGCTCAGTGG 1577
Db 197346 CAATTAATACAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 197405
QY 1578 CAACAAACTTTCTACAGAGACGTTATCTGCTTCTTAACATGTCCAGGCGATGCCAGAC 1637
Db 197406 CAGCAAAATTTTATATGATGAAGATTTACGCAACTAATATGTTAAACGGGAGCCGAAT 197465
QY 1638 ATTAATCTCTCTGTGAAGCTTATGSCATCAAGGTTATTAATCTGTGCGCAAGCGGAGAT 1697
Db 197466 TTTGTTGCTTTAGCTATATGCTTTTGGAGTGGAGGGGTATTAATTAAGAGCCTAATTA 197525
QY 1698 TTGCGCCCGGCGATGCCCGCAATGCTAGCCCAATAGTGTCTGTGTAATGATGATG 1757
Db 197526 TTGAAAGAAAGTTAAAGAAAGCTTTAATTCGAAGAGCAATGCTCATTTGATTTCAAT 197585
QY 1758 GTCAAAAGATGAAAGCTTTTACCTATGATTTGCCCGGATGAGTAAATGCCCAATG 1817
Db 197586 GTCCGTAGAGATGAGATTTGTTATCAATGTGTTCCACAGGTAAAGCAATGCTGAATG 197645
QY 1818 CTAGTTTACCGGA 1831
Db 197646 GTTGTGTTGCTGA 197659

```

Search completed: July 25, 2004, 10:17:16
 Job time : 7875.05 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 25, 2004, 03:57:37 ; Search time 822.718 Seconds

(without alignments)
10115.521 Million cell updates/sec

Title: US-09-893-033-6

Perfect score: 1959

Sequence: 1 gccctaggagccatcgccg.....cataagccaaatgaattc 1959

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: Geneseq_29Jan04:*
2: Geneseqn1980s:*
3: Geneseqn1990s:*
4: Geneseqn2000s:*
5: Geneseqn2001as:*
6: Geneseqn2002s:*
7: Geneseqn2003as:*
8: Geneseqn2003bs:*
9: Geneseqn2003cs:*
10: Geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1958	99.9	1959	6	AAD27607 Synchocyt
2	454.6	23.2	1728	6	ABG90234 M. capsul
3	415	21.2	1728	7	ACA53857 Prokaryot
4	382.6	19.5	1755	7	ACF71672 Photorhab
5	382.6	19.5	110000	7	Continuation (49 o
6	382.6	19.5	110000	7	Act65387 Photorhab
7	381.6	19.5	1878	7	ACA39863 Prokaryot
8	379.6	19.4	1722	7	ACA37690 Prokaryot
9	379.6	19.4	1722	7	ACA53435 Prokaryot
10	378.8	19.3	1875	7	ACA23942 Prokaryot
11	373.8	19.1	1725	4	AA554310 Pseudomon
12	373.8	19.1	1725	7	ACA42854 Prokaryot
13	368.6	18.8	1734	2	AA745893 R. capsul
14	368.6	18.8	1815	4	AA552288 E. coli D
15	368.6	18.8	1815	7	ACA32352 Prokaryot
16	368.2	18.8	1833	7	ACA39024 Prokaryot
17	368.2	18.8	96109	4	AA28548 Genomic f
18	367.8	18.8	1725	7	ACA43064 Prokaryot
19	366.8	18.7	1854	7	ACA38394 Prokaryot
20	366.8	18.7	1857	7	ACA40769 Prokaryot
21	365.8	18.7	110000	4	AA199682_33 o
22	365.4	18.7	110000	4	AA199683_33 o
23	363	18.5	1596	7	ACA44223 Prokaryot

24	363	18.5	1725	6	AA597210	AA597210 Neisseria
25	363	18.5	1725	6	AA597243	AA597243 Neisseria
26	363	18.5	1728	7	ACA41957	ACA41957 Prokaryot
27	358.2	18.3	13732	3	AA81484	AA81484 N. mening
28	358.2	18.3	110000	3	AA81489_1	Continuation (2 of
29	358.2	18.3	110000	3	AA81489_2	Continuation (3 of
30	358.2	18.3	349980	3	AA721611	AA721611 Neisseria
31	355.2	18.1	110000	6	AA803041_20	Continuation (21 o
32	355	18.1	1725	7	AB238433	AB238433 N. gonorr
33	355	18.1	1725	7	ACA41281	ACA41281 Prokaryot
34	354	18.1	1722	7	ACA36592	ACA36592 Prokaryot
35	350.8	17.9	6506	6	ABQ71032	ABQ71032 Listeria
36	349.8	17.9	1761	7	ACA25532	ACA25532 Prokaryot
37	345.8	17.7	1554	7	ACA45757	ACA45757 Prokaryot
38	345	17.6	2286	6	ABQ69168	ABQ69168 Listeria
39	345	17.6	110000	6	ABQ69245_20	Continuation (21 o
40	345	17.6	110000	6	ABQ69245_21	Continuation (22 o
41	345	17.6	319630	6	ABQ67194	ABQ67194 Listeria
42	342.2	17.5	1713	7	ACA27233	ACA27233 Prokaryot
43	338.8	17.3	1725	4	AA500992	AA500992 B. subtil
44	338.8	17.3	3291	4	AA502313	AA502313 B. subtil
45	337.8	17.2	1956	7	ACA29658	ACA29658 Prokaryot

ALIGNMENTS

RESULT 1	AA27607	standard; DNA; 1959 BP.
XX	AA27607	
AC	AA27607;	
XX		
DT	18-APR-2002	(first entry)
XX		
DE	Synechocystis sp. AHA5 large subunit gene, 11vg.	
XX		
KW	Cyanobacteria; acetylhydroxyacid synthase; AHA5; acetoacetate synthase;	
KM	phytoene desaturase; PDS; plant trait control; plasmome transformation;	
KW	herbicide resistance; fungicide resistance; insecticide resistance;	
KM	plant growth protectant; ds.	
XX		
OS	Synechocystis sp.	
XX		
PN	WO200200915-A2.	
PD	03-JAN-2002.	
XX		
PF	27-JUN-2001; 2001WO-US020338.	
XX		
PR	27-JUN-2000; 2000US-0214705P.	
XX		
PA	(AMCY) AMERICAN CYANAMID CO.	
XX		
PI	Kaketuda G, Koop H, Sturmer S, Zhen R;	
XX		
DR	WPI; 2002-139350/18.	
XX		
PT	New cyanobacterial nucleic acid fragments encoding acetylhydroxyacid	
PT	synthase (AHA5) or phytoene desaturase (PDS); useful for conferring	
PT	herbicide, fungicide or insecticide resistance, and for identifying AHA5	
PT	or PDS inhibitors.	
XX		
PS	Claim 1; Page 35-37; 70pp; English.	
XX		
CC	The invention relates to cyanobacterial nucleic acid fragments encoding	
CC	acetoacetate (acetylhydroxyacid) synthase (AHA5) large and small subunits	
CC	or mutant phytoene desaturase (PDS). Polynucleotides of the invention are	
CC	useful for controlling plant traits via nuclear or plasmome	
CC	transformation, e.g. for conferring herbicide, fungicide or insecticide	
CC	resistance. Particularly, they are useful in identifying novel PDS and	
CC	AHA5 inhibitors and in plant transformations for conferring resistance	
CC	and cross-resistance to certain bleaching herbicides and AHA5-inhibiting	

CC herbicides. The present sequence is *Synechocystis* sp. strain PCC 6803
 CC AHAS large subunit gene, *ilvG*
 XX
 SQ Sequence 1959 BP; 465 A; 493 C; 531 G; 469 T; 0 U; 1 Other;
 Query Match 99.9%; Score 1958; DB 6; Length 1959;
 Best local Similarity 100.0%; Pired. No. 0;
 Matches 1959; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCATGAGAGCCCATCCGCGCATGAGTTCATAATTGAGAGCATTAAGCTACGCTTCTTA 60
 DB 1 GCCATGAGAGCCCATCCGCGCATGAGTTCATAATTGAGAGCATTAAGCTACGCTTCTTA 60
 QY 61 ACCGATTGTCAGTGGTTGTCATCAATTCCTAATCCAAAACAAATTTCTGAAAACGTTT 120
 DB 61 ACCGATTGTCAGTGGTTGTCATCAATTCCTAATCCAAAACAAATTTCTGAAAACGTTT 120
 QY 121 CCTTGCCCAACGCGAAACCGGCGCTTATATCTGATGATAGTACCTGAAAGCCATGGGCTC 180
 DB 121 CCTTGCCCAACGCGAAACCGGCGCTTATATCTGATGATAGTACCTGAAAGCCATGGGCTC 180
 QY 181 AAACACATTTTGGCTATCCCGCGGCGCAATTTTGGCCATCTATGATGAACTGTACCGC 240
 DB 181 AAACACATTTTGGCTATCCCGCGGCGCAATTTTGGCCATCTATGATGAACTGTACCGC 240
 QY 241 TTTGAGCGCGCGGCGGAAATTTGACATATTTGGTGGCCATGAAACAGAGCTTCCCAT 300
 DB 241 TTTGAGCGCGCGGCGGAAATTTGACATATTTGGTGGCCATGAAACAGAGCTTCCCAT 300
 QY 301 GCGCGGATGGGATATGCGACAGCCACAGGTAAAGTGGAGTTGTTGGTACATCTGGA 360
 DB 301 GCGCGGATGGGATATGCGACAGCCACAGGTAAAGTGGAGTTGTTGGTACATCTGGA 360
 QY 361 CCAAGGCGGACATACTTGATGACCGGATTTGCCAATGCCATTTGGACTGGTCCCATG 420
 DB 361 CCAAGGCGGACATACTTGATGACCGGATTTGCCAATGCCATTTGGACTGGTCCCATG 420
 QY 421 GTGGATTAATCTGAGAGGTGGCGCGTGCATGATTTGGTAGCGATGTTTCCAGAAATT 480
 DB 421 GTGGATTAATCTGAGAGGTGGCGCGTGCATGATTTGGTAGCGATGTTTCCAGAAATT 480
 QY 481 GACATTTTGGCATCACTTACCGATGCTTAAGCACTCTATGNGATCTAGTGGCGAG 540
 DB 481 GACATTTTGGCATCACTTACCGATGCTTAAGCACTCTATGNGATCTAGTGGCGAG 540
 QY 541 GATATGGCTGGCATTTGTTACTGAGGCTTTTCATCTTGTCTAGCAACGGTCTGTCGGGCG 600
 DB 541 GATATGGCTGGCATTTGTTACTGAGGCTTTTCATCTTGTCTAGCAACGGTCTGTCGGGCG 600
 QY 601 GTTTGATGATATTCCTCAAGAGATGTGGGCTTAGAAGAAATGTAGTACATTCCTCGAC 660
 DB 601 GTTTGATGATATTCCTCAAGAGATGTGGGCTTAGAAGAAATGTAGTACATTCCTCGAC 660
 QY 661 CCGGATGACGTTAATCTACCGGCTTATCGCCCAACGGTTAAAGTATCCCCGACAAATT 720
 DB 661 CCGGATGACGTTAATCTACCGGCTTATCGCCCAACGGTTAAAGTATCCCCGACAAATT 720
 QY 721 AATCGGATGTCATATGTTGAGAGAGGCGAGAAATCCCTGTCTACAGTAGGGGAGGAG 780
 DB 721 AATCGGATGTCATATGTTGAGAGAGGCGAGAAATCCCTGTCTACAGTAGGGGAGGAG 780
 QY 781 GCGATGCGCGCAATGCGCATGCGCCAGAGTGCAGAAATTTGCGGAAAGTTCCAGTTGCCG 840
 DB 781 GCGATGCGCGCAATGCGCATGCGCCAGAGTGCAGAAATTTGCGGAAAGTTCCAGTTGCCG 840
 QY 841 GTAAACAACCACTGATGGGAAATTGGGCGCTTTGAGAGAAACCATCCCTTCGGGGGT 900
 DB 841 GTAAACAACCACTGATGGGAAATTGGGCGCTTTGAGAGAAACCATCCCTTCGGGGGT 900
 QY 901 ATGTGGGATGATGGCCACCGCTATGCCAACTTTGCGGTACAGCAATGTGATTTGTTG 960
 DB 901 ATGTGGGATGATGGCCACCGCTATGCCAACTTTGCGGTACAGCAATGTGATTTGTTG 960

QY 961 ATTGAGATGGGGGCGCGCTTTGAGAGCGGGGTAATGAGCAAACTAGAGCAATTTGCTAGC 1020
 DB 961 ATTGAGATGGGGGCGCGCTTTGAGAGCGGGGTAATGAGCAAACTAGAGCAATTTGCTAGC 1020
 QY 1021 CGCGCAAAAGTAAATTCATTTGACATGACATCGACCGCGGAGGTGGGAAAAACAGGCTTCC 1080
 DB 1021 CGCGCAAAAGTAAATTCATTTGACATGACATCGACCGCGGAGGTGGGAAAAACAGGCTTCC 1080
 QY 1081 GATGTGCCATGTGGGGGAGTGAATGCGCATTTTGAACAGCTTTTGGACGCGGCGCG 1140
 DB 1081 GATGTGCCATGTGGGGGAGTGAATGCGCATTTTGAACAGCTTTTGGACGCGGCGCG 1140
 QY 1141 GAATGTGATTAATCCCAACCATCCCATACAGCCAGCGATGTTAAATGCAATGATCAT 1200
 DB 1141 GAATGTGATTAATCCCAACCATCCCATACAGCCAGCGATGTTAAATGCAATGATCAT 1200
 QY 1201 TGGCGGACCGATTAACCCCTTCCAGGTGCGCCACTATGAGATTAATTTGCCCCCAGAG 1260
 DB 1201 TGGCGGACCGATTAACCCCTTCCAGGTGCGCCACTATGAGATTAATTTGCCCCCAGAG 1260
 QY 1261 GTAGTACACGAAATGTGTGCGCAGGCGCCCGATGCTACTACACACGATGTGGACAA 1320
 DB 1261 GTAGTACACGAAATGTGTGCGCAGGCGCCCGATGCTACTACACACGATGTGGACAA 1320
 QY 1321 CACCAATATGTGGCGCGCCCACTTTTGAACATGCGCCCGCGATGATTTCCAGTGTCT 1380
 DB 1321 CACCAATATGTGGCGCGCCCACTTTTGAACATGCGCCCGCGATGATTTCCAGTGTCT 1380
 QY 1381 GAGCTTGGGATAGATGGGCTTTGTTGTTTACCTGCGCCCATGGAGGCAAAAGTGGAGTGGG 1440
 DB 1381 GAGCTTGGGATAGATGGGCTTTGTTGTTTACCTGCGCCCATGGAGGCAAAAGTGGAGTGGG 1440
 QY 1441 GACGAGCGGTCAATTTGATCAGTGAAGATGCGACCTTCCAAATGATCTTCCAGAACTGG 1500
 DB 1441 GACGAGCGGTCAATTTGATCAGTGAAGATGCGACCTTCCAAATGATCTTCCAGAACTGG 1500
 QY 1501 GAACCTTACCCAGTACGACATCCAGGTTAAATCTATTTCTAATTAACGATGGGAGG 1560
 DB 1501 GAACCTTACCCAGTACGACATCCAGGTTAAATCTATTTCTAATTAACGATGGGAGG 1560
 QY 1561 GGAATGTGCTGCAAGTGCACAACTTCTCTAGAAAGCGTATCTGCTTCTAATCATGT 1620
 DB 1561 GGAATGTGCTGCAAGTGCACAACTTCTCTAGAAAGCGTATCTGCTTCTAATCATGT 1620
 QY 1621 CCCAGGAGTGCACGACATTTAATCTCTCTGTGAAAGCTATGAGTCAAGGGTATTACTG 1680
 DB 1621 CCCAGGAGTGCACGACATTTAATCTCTCTGTGAAAGCTATGAGTCAAGGGTATTACTG 1680
 QY 1681 TGGCAACGCGGAAGATTTGGCGCCCGGAGTCCGCCAAATGCTAGCCCAATGCTCTG 1740
 DB 1681 TGGCAACGCGGAAGATTTGGCGCCCGGAGTCCGCCAAATGCTAGCCCAATGCTCTG 1740
 QY 1741 TGGTATGATGTGTGTGTCAAAAAAGTGAATCTTTAACCCTATGATTTGCCCGGCA 1800
 DB 1741 TGGTATGATGTGTGTGTCAAAAAAGTGAATCTTTAACCCTATGATTTGCCCGGCA 1800
 QY 1801 TGAATGATGTCCTAATGTAGTGTATACCGGAAGTGGCGGATGACGAGTACNGAGCAATGGTCCCGGA 1860
 DB 1801 TGAATGATGTCCTAATGTAGTGTGTATACCGGAAGTGGCGGATGACGAGTACNGAGCAATGGTCCCGGA 1860
 QY 1861 TGGTGAATGTCACCAATTTGCCAAACCCCAAAATTTTCAATCACCAATGCTTGTCTG 1920
 DB 1861 TGGTGAATGTCACCAATTTGCCAAACCCCAAAATTTTCAATCACCAATGCTTGTCTG 1920
 QY 1921 GTGAGGCCAAATCTTAACCCCAATTAAGCCAAATTTGAATTC 1959
 DB 1921 GTGAGGCCAAATCTTAACCCCAATTAAGCCAAATTTGAATTC 1959

RESULT 2
 ABO90234
 ID ABO90234 standard; DNA; 1728 BP.
 XX

AC AB090234;
XX
DT 01-OCT-2002 (first entry)
XX
DE M. capsulatus gene #219 for DNA array.
XX
KM Micro array; gene; ds; differential expression; gene expression.
XX
OS Methyllococcus capsulatus.
XX
PN W0200255655-A2.
XX
PD 18-JUL-2002.
XX
PF 14-JAN-2002; 2002WO-N0000019.
XX
PR 12-JAN-2001; 2001NO-00000235.
XX
PR 12-JAN-2001; 2001NO-00000239.
XX
PA (UNIF-) UNIFOB STIFTELSEN UNIV BERGEN.
XX
PI (TIGR-) TIGR.
XX
PI Birkeland NK, Eidhammer I, Jonassen I, Jensen HB, Lien T;
PI Lillehaug JR, Lossius I, Eisen JA, Fraser CM, Durkin AS;
PI Salzberg SL;
XX
DR WPI: 2002-557818/59.
XX
XX
XX Novel DNA array useful for determining differential expression of
PT Methyllococcus capsulatus genes, comprises polynucleotides or
PT oligonucleotides representative for a selective number of Methyllococcus
PI capsulatus genes.
XX
XX
XX Claim 19; Page 156-157; 678BP; English.
XX
XX
XX The invention relates to a novel DNA array giving a representation of a
CC number of Methyllococcus capsulatus genes. The method of the invention is
CC useful for determination of the differential expression of the genes of
CC M. capsulatus, and for studying gene expression on a genomic scale and in
CC gene expression assays of M. capsulatus genes. The sequences shown in
CC AB090016-AB091855 represent M. capsulatus genes for use in arrays of the
CC invention
XX
SQ Sequence 1728 BP; 356 A; 576 C; 506 G; 290 T; 0 U; 0 Other;

Query Match 23.2%; Score 454.6; DB 6; Length 1728;
Best Local Similarity 56.8%; Pred. No. 2.6e-133;
Matches 962; Conservative 0; Mismatches 709; Indels 22; Gaps 6;

QY 136 ACCGGGCTTATATCTGATGATAGCTGAAACCGCATGGGGTCAACACATTTTGGC 195
DB 10 AGCGGTGGGAAATGCTGTCCTTAAGACGAGGGTGTGATATCATCTTGGGC 69
QY 136 TATCCCGGGGGCAATTTTGGCCCATCTATGATGAAGTACCGCTTTGAAGCGGGG 255
DB 70 TATCCGGGGGGTGAAGTATGATATTAAGATGGCTGTTCAAACAGAGAGATG----- 124
QY 256 GAAATTTAGCATATTTTGGTGGCGCATGAACAAGAGCTTCCATGGCGGAGATGGTAT 315
DB 125 -----TAAAGCATCTCTGTGTGGCCACGAACGAGGCGGACCCGACGCGGATAC 180
QY 316 GCCAGAGCCACAGTAAAGTGGAGATTGTTTGGTATCATCTGACACAGGGGCGACTAAC 375
DB 181 GCCCCTCCACCGGAAACCCCGCGTGTCTGTGACTTGGGGCCCGGTCGACCAAT 240
QY 376 TTGGTGAACCGGATGGCAATGCGCATTTGACCTGGGTGCCATGGTGGTATTAATGGA 435
DB 241 GCGGTACCGGATGCGCACCGGCTTACATGATTTGATTCGCGTGTGTGATCAACGAGC 300
QY 436 GAGGTGGGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 495
DB 301 CAGGTGGGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360

QY 496 ACCTTACCATGCTTTAAGACATCTCTATGTGTGATGATGATGATGATGATGATGATGAT 555
DB 361 ACTGGGCGGTGCTGAAAGCAAACTTCTGTGTCAAGAGATCATCAAGCTGGCCGAAACT 420
QY 556 GTTACTGAGGCTTTTCAATCTTTGTAGACCGGTGTCGCGGCGGCTTTTGTATGATTT 615
DB 421 TTAAAGAGGCTTTTCAATCTTTGTAGACCGGTGTCGCGGCGGCTTTTGTATGATTT 480
QY 616 CCCAAGATGTGGGCTTTAAGATGTGATGATGATGATGATGATGATGATGATGATGATGAT 675
DB 481 CCCAAGATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
QY 676 CTACCGGCTTATGCGCCACCGTTAAAGTATCCCGCAAAATTAATGCGGATTTGCA 735
DB 541 CTAGGCTTCAATCTTTGTGGAAGGGGATCCCATCAATCAATCAATCAATCAATCAAT 600
QY 736 TTGTGAGAGGCGCAAAATCCCTGTCTCAATGAGGGGAGAGGGGAGATGCGGCGCAAT 795
DB 601 CTATGCTGTGCGCTGCGCGCGCGATGATCTACAGCGGCGGCGGCTATCTCGGCAAT 660
QY 796 GCCCATGCGGATGCGAGATTTGCGGAAAGTTCCAGTTGCGGTAACAACACCTTG 855
DB 661 GCGCGGACGAACTGACCAAGCTGACGCGGCTGCTGATCCCATCAACCAACCTTG 720
QY 856 ATGGGAATTTGGGCTTTTGAACGAACACCCCTTGGTGGGATGTTGGGTATGAT 915
DB 721 ATGGGTCTGGGCGGCTTACCCCGCACCAACCGGCTGCTGCTGCTGCTGCTGCTGCTG 780
QY 916 GGCACCGCTATGCAATCTTTGCGTCAAGCATGATGATGATGATGATGATGATGATGATGAT 975
DB 781 GGCACCTAGAGGCAACATGAGGATGACGATGACGATGACGATGACGATGACGATGACGAT 840
QY 976 CGTTTGAAGACCGGCTTATGCGCAACTGAGAGATTTGCTAGCGGCGGCAAGTAT 1035
DB 841 CGCTTGAAGACCGGCTTATGCGCAACTGAGAGATTTGCTAGCGGCGGCAAGTAT 900
QY 1036 CACATTGACATGACCGCGGAGGTGGGAAAAAACAAGGCTTCCCATGTCCTCATTTG 1095
DB 901 CACATGACGATGACCGCGGCTGATCTCCAAACCGTGAAGTGAACATCCCATCTGTC 960
QY 1096 GGGGATGTACGATGATTTTGAACAGCTTTTGAACAGCTTTTGAACAGCTTTTGAACAG 1155
DB 961 GGGGATGTACGATGATTTTGAACAGCTTTTGAACAGCTTTTGAACAGCTTTTGAACAG 1020
QY 1156 ACCATGCCATATCAACCGGATGATGATGATGATGATGATGATGATGATGATGATGAT 1215
DB 1021 CCGATGCCGAGGCGCTGAAAGGCTTGTGGCAAGATGATGATGATGATGATGATGATGAT 1080
QY 1216 CCGCTCCAGTGGCCCATATGAGATGATGATGATGATGATGATGATGATGATGATGAT 1275
DB 1081 TGCTTGGCTACGACCGGCAAGCAAGATGATGATGATGATGATGATGATGATGATGATGAT 1140
QY 1276 GTTGGGAGGCGGCGGCTTGTGAGAGCTTTTGAACAGCTTTTGAACAGCTTTTGAACAG 1332
DB 1141 TGGGAATGACGCGGCGGCGGCTTGTGAGAGCTTTTGAACAGCTTTTGAACAGCTTT 1200
QY 1333 GCGGCGGCTTTTGAACGCTTGTGAGAGCTTTTGAACAGCTTTTGAACAGCTTTTGAACAG 1389
DB 1201 GCGGCGGCTTTTGAACGCTTGTGAGAGCTTTTGAACAGCTTTTGAACAGCTTTTGAACAG 1260
QY 1390 ACCATGGGCTTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1448
DB 1261 ACCATGGGCTTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320
QY 1449 GTCATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1508
DB 1321 GTCATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1380
QY 1509 GCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1568
DB 1381 TTGCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1440
QY 1569 CGTACGTGCAAAACATTTTCTACGAAGAACGTTATTTCTGCTTCTAATCATGTCAGGGC 1628

Db 1441 CGACATGCGAGGATTCATCATGAAAGCCGCTATTCGATCTCTACTG---GAGACT 1497
 Qy 1629 ATGCCAGCATTAATCTCTGTGGAAGCCATATGGAATCAAGGATATTAATGTCGGCAG 1688
 Db 1498 CTGCGGACTTCTCAAACTGGCGAGGCTTACGTCACGTGGCAGTCCGATGACAGA 1557
 Qy 1689 CGGGAAGATTGGCCCGCGCATGCCGAAATGCTAGCCCAATG---GTCTGTGTG 1745
 Db 1558 CCGGCGGAGCTGGCCCGCCCTGGAAGAGCCCTGAAACTCAAGACCGACCGGTCTTC 1617
 Qy 1746 ATGATGTGTGTGTCAAAAGATGAAGAACTGTATACCTTATGATGCCCCGGATGAGT 1805
 Db 1618 ATGATTTCTTCCACCGATCCCAACGAAAGCTTACCCCATGTGATCCGGAAGGCC 1677
 Qy 1806 AATGCCCAATGC 1818
 Db 1678 CATCACGAGATGC 1690

RESULT 3
 ID ACAS3857 standard; DNA; 1728 BP.
 AC ACAS3857;
 XX ACAS3857;
 XX 19-JUN-2003 (first entry)
 DE Prokaryotic essential gene #35514.
 KM Antisense; ds; prokaryotic essential gene; cell proliferation;
 KM drug design; gene.
 OS Versinia pestis.
 XX WO20027183-A2.
 PD 03-OCT-2002.
 XX 21-MAR-2002; 2002WO-US009107.
 PF 21-MAR-2001; 2001US-00815242.
 PR 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362639P.
 XX (ELIT-) ELITRA PHARM INC.
 XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 XX WPI; 2003-029926/02.
 DR P-PSDB; ABU4987.
 XX New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 PS Claim 14; SEQ ID NO 41727; 1766bp; English.

The invention relates to an isolated nucleic acid comprising any one of the 6213 antisense sequences given in the specification where expression of the nucleic acid inhibits proliferation of a cell. Also included are: (1) a vector comprising a promoter operably linked to the nucleic acid encoding a polypeptide whose expression is inhibited by the antisense nucleic acid; (2) a host cell containing the vector; (3) an isolated polypeptide or its fragment whose expression is inhibited by the antisense nucleic acid; (4) an antibody capable of specifically binding the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular proliferation or the activity of a gene in an operon required for proliferation; (7) identifying a compound that influences the activity of the gene product or that has an activity against a biological pathway

CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a gene
 CC product's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target
 CC prokaryotic essential genes. Note: The sequence data for this patent did
 CC not form part of the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pat_sequences

XX Sequence 1728 BP; 416 A; 387 C; 473 G; 452 T; 0 U; 0 Other;
 SQ

Query Match 21.2%; Score 415; DB 7; Length 1728;
 Best Local Similarity 55.8%; Pred. No. 1,1e-120;
 Matches 929; Conservative 0; Mismatches 705; Indels 31; Gaps 6;

Qy 172 CATGGGCTCAACACATTTTGGCTATCCCGCGGGCAATTTGCCATATGATCA 231
 Db 49 CAGGGCGTTAAGACATGATTCGGTTATCCCGCGGGCGGATGATTTAAGATGCC 108
 Qy 232 CTGTACCGCTTTGAAGCGCGGGGAATGAGCATATTTTGGCGCCATGAACAAGA 291
 Db 109 CTGCAC-----ACGGTGGAGGCACTGATCAGCTGTGTGGCGCAAGAACAGT 159
 Qy 292 GCTTCCCATGCGCGGATGGATGATCCAGACCAAGGTAAAGTGGAGTTTGTGGT 351
 Db 160 GCGGTTCAATGCGCGATGATGATGCGCGGCAACGCGCGATGATGATGATGATG 219
 Qy 352 ACATGTGACAGCGGGGCACTAATCTGTGACCGGCAATGCAATGCCATTTGAGCTCG 411
 Db 220 AGCTGTGCTCTGTGCAACCAATGCGATTCACGGTATGCGACCGCTTATATGATTA 279
 Qy 412 GTGCCATGTGTGATTAATCTGAGAGGTGGCCCGTGCATGATTTGGATGCGATCTTC 471
 Db 280 GTGCCAATGTGTGCTCTTCAAGCAAGTGCACAGCTATGATTTGGTTGATGCTTC 339
 Qy 472 CAGGAATTTGACATTTTGGGATCACTTACCGATCGTTAGACCTCTATGTGTAGCT 531
 Db 340 CAAGATGTGACATGTGGGATCTCCCGCAGTGTCAACATAGCTTCTTGTATAG 399
 Qy 532 AGTGGCGGATATGATGCTCGCATTTTACTGAGCTTTCATCTTGTAGCACCGGTCTG 591
 Db 400 CGCAAGAAACATCCGATGATGTGTGAAAAAGCCTTTTATCTGGGATCCACAGCGCT 459
 Qy 592 CCGGCGCGGTTTGTATGATTAATCCCAAGATGTGGCTTAAAGAAATGTGATCAATT 651
 Db 460 CCGGCTCGGTGTATGATGACTTGGCAAAACATTTGTGCGCCACGAGATTAATGCT 519
 Qy 652 CCGCTGACCGCGGAGCTTAATCTACCGGTTATGCGCCACGAGTTAAAGTATGCC 711
 Db 520 TAGGCTTATCCAGAGCAAGTCAAGTTGGCTCTCTATATCTACGTGCAAGGACCGCT 579
 Qy 712 CGACAAATTAATGCGGATTCGAATTTGTGAGAGCGCCAGAAATCCCTTGTCTAGTA 771
 Db 580 GGGCAGATTAATGCGGATTCGAATTTGTGAGAGCGCCAGAAATCCCTTGTCTAGTA 771
 Qy 772 GGGGAGGGGGATGCGCGCAATGCCCATGCCAGGTGACGAATTTGGGAAAGGTTC 831
 Db 640 GGTGTGTGTGATTAATGCGGCTTGTCCATTAAGAAATTTACTGATTTGGCAGAAAGT 699
 Qy 832 CAGTTGCGGATTAACCAACCGCTGATGGAATTTGGGCTTTTGAAGAAACATCCCTT 891
 Db 700 AACTTACCGATTAACAGGCTCTGTATGGGCTTAGCAGTTTCCTGTAGCATGCGCAA 759

OY		892	TGCGGGGTATGTTGGGATGCATGTCACCCGCTATGGCAACTTGGCCGTAGCGAANTG	951
Db		760	AGTGTCGGCAGCTTGGGGATGCAATGTGACTTTGAAGCCAAATATGBCCATGACCAATACT	819
OY		952	GATTGTTGATTGCAATGGGGGCCCGTTTCGACGACGGGGTAATGCGAAACTAGACGAA	1011H
Db		820	GATCTGATTTTTGGCCGTGGGTGTCGGTTTACGATCGTACCACCAATAATTTTAGCAAA	879
OY		1012	TTTGCTAGCCCGCCAAAGTAATTCATTCAGATTCGACCCGGGGGAGAGTGGAAAAAC	1071H
Db		880	TATTCGCCCTGATGCCACAGCTGTCACTCATTTGATATCGATCCGACATCGATCTCGAAACG	939
OY		1072	AGGCGTCCCGAGTGGCCCATGTGGGGGATGACGCATGTTTTAGAACAGCTTTTGCA	1131H
Db		940	GTCACCCGCGCATATCCCAGATTGGGGTATGCCAAACAAGTAACTGACACAAATGTTAGAT	999
OY		1132	CGGGCCCGGGGATTTGATTTACCCACCCATCCCATACCCAGCAGATGTTAATATGC	1191H
Db		1000	TTGTTGTGCAAGTGTGGCTGTACACAGAATTTTGTATGCTTGGCGGATGTGGGCATTC	1059S
OY		1192	ATTGATCATTTGGCGGACCGATTATCCCCCTCCAGGTGCCCATATGAGCATACTATTGGC	1251H
Db		1060	ATTGAAACATGGCGCACCCGTAATTTGCTTGGGCTACAAACAAAGACATGGGAAAAATCAA	1119
OY		1252	CCCCAGAGGTATGACACGAATTTGTGTCCAGGCCCC--CGATGCTACTACACACC	1308S
Db		1120	CCACAAGGGGTGATTGAAGCCCTTCATGTGTGATTAAGGCGATGCGTACGTTACCTCT	1179H
OY		1309	GATGTGGGACAACAACAATGTGGCGGGCCCC--AGTTTGAACAATATGCCCCCGCGGA	1365S
Db		1180	GATGTGGGCCAACATCAGATGTTTGGGCTCTTAACTATCCCTTTGATTAACCAAGTGTGT	1239S
OY		1366	TGGAATTCAGATGCTGGCTTTGGGTACGATGGGCTTTGGTTAACTGCGCCCATGGAGCC	1425S
Db		1240	TGGATTTAACTCCGGTGGCTTGGGCAAGATGGGCTTTGGCTTGCACGCGCCCTTGGGGTT	1299H
OY		1426	AAAGTGGAGTGGGGGAGCG-AGCGGTCAATTTGCATCAGTGGAGATGCGCAGCTCCAATG	1484H
Db		1300	AAATTTGGCTTTGCCMAATGAAACCGCTCTTTGTGTGACCGGTGATGGCAGTATCCAGATG	1355S
OY		1485	AATCTTCAGAACTGGGGAACCTTAGCCAGTACGACATCCAGGTTAAACTATTAATTTCTC	1544H
Db		1360	AAATATTCAGGAGTTATCAACCGCACTGCAMAATAACTTGCCACAGTGTGTATTAACCTG	1419
OY		1545	AATAACGGTTGSCAGGGGATGGTGTGAGTGGGCAACAACCTTCTACGAAGAACGTTAT	1604H
Db		1420	AATATACCCCTATCTGGGGATGTGTAACAAATGGCAGCAGATTTATTATTCAGGCGCGCAT	1479H
OY		1605	TCTGCTCTTCAATGTCCTCCAGGGCATGCGACACATTAATCTCTCTGTGAAGCCTATGSC	1664H
Db		1480	TCTCAGTGTGATATG---GATTTGCGTGCCTGATTTTGTCAAGTTGGCTGGAAGCCTATGSC	1536H
OY		1665	ATCAGAGGTATTAATCTGTGCGCAAGCGGGGAAATTTGGCCCGGGGAGTCCGGAATGTA	1724H
Db		1537	CATATTGGTGTGGCTAATTAATACGCGGGATAGCTGGAAGTAAAGCTGGCAGACGCACTG	1596H
OY		1725	GCCCA-----CAATGTCCTGTGTGTGATGATGTGTGTGTCAAAAAAGATGAA	1772H
Db		1597	ACACAATTAATCCGAAACCAATCGTTTGGTCTTTTGGATGTGATGCTGATGAACAAGAA	1656H
OY		1773	AACGTTTAACCTATGATTTGGCCCGGACATAGATATGCCCAATG	1817
Db		1657	CATGTTTACCGGATGCAGATTTGTGTGTGTGTGGCATGACGAAATG	1701

RESULT 4	
ACF71672	
XX	ACF71672 standard; DNA; 1755 BP
AC	ACF71672;
XX	
DT	20-NOV-2003 (first entry)

DE Photorhabdus luminescens nucleotide sequence #10139.

KM Antibacterial; fungicide; insecticide; polymorphism; genetic analysis;

KM detection; food; gene expression; plant; animal; microorganism; toxin;

KM antibiotic; bioplastic; virulence factor; disease model; plague;

KM whooping cough; gene; ds.

OS Photorhabdus luminescens.

XX NC0200294867.A2.

XX

XX 28-NOV-2002.

PD

XX 07-FEB-2002; 2002WC-IB003040.

PF 07-FEB-2001; 2001R-00001659.

XX

PR (INSP) INST PASTEUR.

FA (CNRS) CNRS CENT NAT RECH SCI.

PA

XX Duchaud E, Taourit S, Glaser P, Frangeul L, Kunst F, Danchin A,

PI Buchrieser C;

XX WPI; 2003-148459/14.

XX

PT Genomic sequence of Photorhabdus luminescens and encoded polypeptides,

PT useful e.g. as therapeutic antimicrobials and agricultural pesticides.

PS Claim 2; SEQ ID NO 10139; 1205bp; French.

XX

XX The invention relates to the isolation of genes and their encoded

CC proteins from Photorhabdus luminescens. The isolated sequences are

CC sources of probes and primers for detecting the genome of P. luminescens

CC and related species; to study polymorphisms; for gene analysis and for

CC detection/amplification of the genes. Antibodies (Ab) raised against the

CC polypeptides encoded by the genes are used for detection/identification

CC of P. luminescens; e.g. in foods. The genes, proteins, Ab and cells that

CC carry a gene-containing vector are used to select compounds that

CC modulate, regulate, induce or inhibit expression of the genes in plants,

CC animals or microorganisms other than P. luminescens and are able to alter

CC response or sensitivity to toxins and antibiotics produced by P.

CC luminescent. Cells transformed to express the genes are useful for

CC recombinant production of the proteins, particularly toxins and

CC antibacterials useful as insecticides, bactericides and fungicides. The

CC genes, proteins, vectors containing the genes and Ab are also useful

CC therapeutically (to treat microbial infection by bacteria or fungi that

CC are sensitive to P. luminescens-encoded toxins or antibiotics) and as

CC bioplastics. Other uses of the genes and the proteins are as virulence

CC factors and for identifying targets of human diseases for which P.

CC luminescens is a model (particularly plague and whooping cough). This

CC sequence represents one of the isolated P. luminescens genes

XX

SQ Sequence 1755 BP; 484 A; 322 C; 459 G; 490 T; 0 U; 0 Other;

Query Match 19.5%; Score 382.6; DB 7; Length 1755;
Best Local Similarity 54.6%; Pred. No. 2.2e-110;
Matches 877; Conservative 0; Mismatches 709; Indels 19; Gaps 5

Dy 121 CCTAGCAACGGCAAACCGGCGCTTATTCCTGATGATGACCTGAAACGCCATGGGAGTC 180
| | | | |
Db 25 CCTATGAAGATGTATTACAGAGCTGAATGCTGCATCGATCGTTAATTATGACGAGCGATT 84
| | | | |
Dy 181 AAAACAATTTTGCGTATCCCGCGCGGCAATTTGCCCATCTATGATAACTGTATCCGC 240
| | | | |
Db 85 AACACATATTTCTGTATTCGCGGGCGGCACTAGTATTTATGATGCGCCGTCAAT--- 141
| | | | |
Dy 241 TTGTAAGCGCGCGCGGAATTTAGACATATTTTGTTGCGCATATAACAAGAGCTTCCCAT 300
| | | | |
Db 142 -----ACGGCTGGGGGCAATTGAGATATCTCGTAGTATAGACAGCGGCGCAGTTCA 195
| | | | |
Dy 301 GGCGCGAATGGGTATGCAAGCAACGTAAGTAAAGTGGAGTTGTTTGGTACATCTGGA 360
| | | | |

Db 196 ATGCTGATGTTATGCTCGATCTACCGGGAAGTTGTTGATTTAATCACCCTAAGT 255
 Qy 361 CCAGGGGCGACTAACTTGGTACCGGCGATTGCCAATGCCATTGGACTCGTGCCCAT 420
 Db 256 CCAGAGCAACCAATGCCATTACCGGATTGCCACAGCTTATATGATTCATTCGATG 315
 Qy 421 GTGGTATTTACTGAGAGGTGGGCGCCGTCATGATTTGGTACGATGTTTCCAGGAAT 480
 Db 316 GTTGTATTTGCTGGTCAAGTCGACGATTCGTTATGATGATGATTCATTCAGAAAT 375
 Qy 481 GACATTTTGGCATCACCTTACCGATCGTAAAGCACTCTATGTTGATGATGCGGCG 540
 Db 376 GACATGTAAGGATATCCCGCCGATCGTAAACAGTTTATGAGAAAAAAGCAGAG 435
 Qy 541 GATATGCTCGCATTTGTTACTGAGCTTCCATTTGCTAGCAACCGGTGTCCTCGGCG 600
 Db 436 GATATTCGAAAGTACTGAAAAAGCTTTCTATCTACTTCCAGTGTCTCGGGAACC 495
 Qy 601 GTTTGATGATATTTCCCAAGATGATGAGTGAAGATGATGATGATGATGATGATG 660
 Db 496 GTGCTTATGATTTACCAAAAAGACAGGTAAATCCGCACTAAATTAACCTATATTTAT 555
 Qy 661 CCGGCTGACGTTAATCTACCGGATATCGCCCAACGCTTAAAGTAAATCCCGACAAAT 720
 Db 556 CCTGATCAGATTAAGTATGCGTTCTTATATTCACAGATTCAGAGACCAAAAGGCAAGAT 615
 Qy 721 AATGCGCATTTGCATTTGTTGAGACAGGCGAGAAATCCCTGCTCTACGAGGAGGAG 780
 Db 616 AAGCGGCGATTAAAGAACTATGATGCGAAAAACCACTCATATATGAGCGGCGG 675
 Qy 781 GCGATCCCGCCCAATGCCATGCCAGTGCAGGAATTTGCGAAAGTTTCCAGTTGCCG 840
 Db 676 GCGATTAACGCAATGTTGATGAGCAAAATTAAGTATGAGCAAAATTAATTCCT 735
 Qy 841 GTAAACAACCAACCGATGAGGAATTTGGGCGTTTGAAGCAAAACATCCCTTGCGTGGG 900
 Db 736 GTTGTAGCTCCCTGATGAGGCTTAGGCAATTTCCCTGCACTCATCCGCAACCTGGG 795
 Qy 901 ATGTTGGATATGATGCGCAACCGTATGCAACTTTGCGTACGAGATGATTTGTTG 960
 Db 796 ATGCTCGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 855
 Qy 961 ATTGCAATGAGGCGCGCTTTCAGACGCGGTAATGCGCAACTAGACGATTTGCTAG 1020
 Db 856 TTTGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 915
 Qy 1021 CGCGCCAAAGTATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 1080
 Db 916 GAACCAACGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 975
 Qy 1081 GATGTCCTCATTTGGGAGATGATGCGCATGTTTGAACAAGCTTTTGCAGCGGCGCC 1140
 Db 976 GATGATCCTATTTGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1035
 Qy 1141 GAATTTGATTTCCCAACCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 1200
 Db 1036 ACACTAAGAC 1095
 Qy 1201 TGGCGAGCCGATTAATTTAGTTGATGATGATGATGATGATGATGATGATGATGAT 1260
 Db 1096 TGGCGAGCCGATTAATTTAGTTGATGATGATGATGATGATGATGATGATGATGAT 1155
 Qy 1261 GTAGTACACGAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1317
 Db 1156 GCTTTCAGGTTCTCTATGCTTACAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAG 1215
 Qy 1318 CAACACCAAAATGTTGGGCGCCAGATTTTGAACAATGG---CCCCCGCGATGATTTCC 1374
 Db 1216 CAGCATCAGATGTTGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1275
 Qy 1375 AGTGTGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1433
 Db 1276 TCCGTTGTTGGGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1335

Qy 1434 ACTGGGGAGACAGCCGCTATTTGCATGATGAGATGCCAGCTTCCAAATGATCTTCA 1493
 Db 1336 AAACCGAAGCGACGATGCTGTGTTACGGGGATGGTATATTCAGATGAATATTCAG 1395
 Qy 1494 GAATGGAACCCCTAGCCAGTACGACATCCAGGTTAAACTATTTATTCATTAACGAT 1553
 Db 1396 GAGCTGTCCACCGCATTTGCATATGATGATGATGATGATGATGATGATGATGATGAT 1455
 Qy 1554 TGGCAGGGAGATGTTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1613
 Db 1456 TTTCTGGGAGATGATGAAGCAATGACATGATGATGATGATGATGATGATGATGATGAT 1515
 Qy 1614 AACATGCTCCAGGCGATCGACATATATCTCTGTGTAAGCTTATGATGATGATGAT 1673
 Db 1516 TATATG---GATCTTCTGCTGATTTTGTCAAAATGCTGATGATGATGATGATGAT 1572
 Qy 1674 ATTACTGTGCGCAAGCGGGAAGATTTGCGCCGCGATTCGCGAA 1718
 Db 1573 ATTCTATTCACGTTATGATGAATGGAATGGAATGGAATGGAATGGAATGGAAT 1617

RESULT 5
 ACF67367_48
 Continuation (49 of 57) of ACF67367 from base 4800001 (Photobabidus luminescens nuc1eot1
 WP Sequence split into 57 fragments LOCUS ACF67367 Accession ACF67367
 WP Fragment Name Begin End
 WP ACF67367_01 1 11000
 WP ACF67367_02 20001 31000
 WP ACF67367_03 30001 41000
 WP ACF67367_04 40001 51000
 WP ACF67367_05 50001 61000
 WP ACF67367_06 60001 71000
 WP ACF67367_07 70001 81000
 WP ACF67367_08 80001 91000
 WP ACF67367_09 90001 101000
 WP ACF67367_10 100001 111000
 WP ACF67367_11 110001 121000
 WP ACF67367_12 120001 131000
 WP ACF67367_13 130001 141000
 WP ACF67367_14 140001 151000
 WP ACF67367_15 150001 161000
 WP ACF67367_16 160001 171000
 WP ACF67367_17 170001 181000
 WP ACF67367_18 180001 191000
 WP ACF67367_19 190001 201000
 WP ACF67367_20 200001 211000
 WP ACF67367_21 210001 221000
 WP ACF67367_22 220001 231000
 WP ACF67367_23 230001 241000
 WP ACF67367_24 240001 251000
 WP ACF67367_25 250001 261000
 WP ACF67367_26 260001 271000
 WP ACF67367_27 270001 281000
 WP ACF67367_28 280001 291000
 WP ACF67367_29 290001 301000
 WP ACF67367_30 300001 311000
 WP ACF67367_31 310001 321000
 WP ACF67367_32 320001 331000
 WP ACF67367_33 330001 341000
 WP ACF67367_34 340001 351000
 WP ACF67367_35 350001 361000
 WP ACF67367_36 360001 371000
 WP ACF67367_37 370001 381000
 WP ACF67367_38 380001 391000
 WP ACF67367_39 390001 401000
 WP ACF67367_40 400001 411000
 WP ACF67367_41 410001 421000
 WP ACF67367_42 420001 431000
 WP ACF67367_43 430001 441000
 WP ACF67367_44 440001 451000
 WP ACF67367_45 450001 461000

WP ACF67367_46 460001 4710000
 WP ACF67367_47 470001 4810000
 WP ACF67367_48 480001 4910000
 WP ACF67367_49 490001 5010000
 WP ACF67367_50 500001 5110000
 WP ACF67367_51 510001 5210000
 WP ACF67367_52 520001 5310000
 WP ACF67367_53 530001 5410000
 WP ACF67367_54 540001 5510000
 WP ACF67367_55 550001 5610000
 WP ACF67367_56 560001 5648994

Query Match 19.5% Score 382.6; DB 7; Length 110000;
 Best Local Similarity 54.6% Pred. No. 2.5e-109;
 Matches 877; Conservative 0; Mismatches 709; Indels 19; Gaps 5;

QY 121 CCTAGCAACGGCAACCGGCGCTTATATCTGATGATGAGCTGAAACGCCATGGGCTC 180
 DB 66022 CCTATGAGATGTATACAGAGCTGAATGTCGTCCGATGCTTATGATGATGAGGCGCTT 66081
 QY 181 AAACACATTTTGGGCTATCCCGCGGGGCAATTTTGCCCATCTATGATGAAGTGTACCGC 240
 DB 66082 AAGCATATATCTGTTATCCGGGGGGGCGAGTACTGATATTTATGATGATGCCCTGCAT--- 66138
 QY 241 TTGGAAGCGCGGGGGAATTGAGCATATTTTGTGCGCATGAAACAAGAGCTTCCCAT 300
 DB 66139 -----ACGGCTGGGGGCGATTGAGCATATCCCTGATGCTGATGAGCAAGGGGCGAGTCAAT 66192
 QY 301 GGGGCGATGGGTATGCCAGAGCCACAGTAAAGTGGAGTTTGTTCGTACATCTGGA 360
 DB 66193 ATGGCTGATGTTATGCTGATGATCTACCGGGAAGTTGGTGTGTTATATCACTACAGT 66252
 QY 361 CCAGGGGCGACTACTACTGCTGACCGGCAATGCGCAATGCCAATGCCAATTGCACTCGGTCGCA 420
 DB 66253 CCAGAGCAACCAACAGCCATTACCGGATTTGCCACAGCTTATATGATGATCAATTCGATG 66312
 QY 421 GTGTGATTAATCTGAGAGGTGGCGGTGCCATGATTTGGTATGAGATGCTTTCCAGAAAT 480
 DB 66313 GTTGTATTTGTCGGTCAAGTCCAGTCCAGTTCGTTGATCGGTAATGATGATTCAGAAATGT 66372
 QY 481 GACATTTTGGGATCACTTACCGATCGTTAAGCATCTCTATGTGTACGTATGTGCGCG 540
 DB 66373 GACATGTAGGATATCCCGCCGATCGTAAACACAGTTTATGTAATAAAAGCAAG 66432
 QY 541 GATATGGCTCGCATTTGTAAGGCTTTCCATCTTGTAGACACCGGCTGTCGCGGCG 600
 DB 66433 GATATTCGAGAGTACTGAAAAAAGCTTTCTATCTAGCTTCCAGTGTGCTCCGGAACC 66492
 QY 601 GTTTGATGATATTCCTCAAGGATGTGGGCTTGAAGAATGTGATCAATTCCTCGAC 660
 DB 66493 GTGCTATTGATTTTCAAAAGACACGTAATCCGGCAGTAATAATTACCTATATTTAT 66552
 QY 661 CCGGCTGACGTTATCTACCGGCTATCGCCCAAGCTTAAAGTAATCCCGCAAAAT 720
 DB 66553 CCTGATACGATATGATGCTTCTTATATCAAGATTCAGAGCAACAAAGGCGAGATT 66612
 QY 721 AATGCGGCAATTTGCAATTTGTGAGAGCGCAAAATCCCTTCTCTACGTAGGGGAGAGG 780
 DB 66613 AAGCGGCAATTTAAGAAACTATTTGAATGCAAAAAACAGTATATGTAGGCGGGGG 66672
 QY 781 GCGATCGCGCAATGGCCCATGCGAGGTGAGAGAAATTTGGGAAAGTTCCAGTGTGCG 840
 DB 66673 GCGATTAACGAGATTTGTCAGCAGAAATTAAGTGTGGGAAAAAATTAATATCTCT 66732
 QY 841 GTACACACACCTGATGAGGAATTTGGGCTTTTTCAGAAAAACATCCCTCTCGGTGAGT 900
 DB 66733 GTTGTAGCTCCCTGATGGGCTTAGGCGAGTTTCCCTGCACTATCCGCAAAAGCTGGG 66792
 QY 901 ATGTTGGGTATGATGCGCACCGCTATGCCAATTTTGGCGTACAGGAATGTGATTTGTG 960
 DB 66793 ATGCTGGGTATGATGATCAATTTGAAAGCAAAAGGCAATTCATTAATTCGATGTCAAT 66852
 QY 961 ATTGCAATGGGGCCCGTTTTCAGCAACCGGGTAACTGGCAAACTAGACGAATTTGTAGC 1020

DB 66853 TTTCTGTGCGGTGATGCTTTTGTGATGATGATGATCAACCAATATCTGAGAAATATTTGCCG 66912
 QY 1021 CGGCCCAAGTAATTCATTTGATCATGACATCCCGCGGAGGTGGGAAAAACAGGCTCCC 1080
 DB 66913 GAACCAACGATTAATTAATTTGATGATGATGATGATCAATCAATTTGAAAAAGTTTCCGA 66972
 QY 1081 GATGTGCCAATTTGGGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
 DB 66973 GATGTACTATTTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 67032
 QY 1141 GAATTTGATTAACCCCAACCCATCCCATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
 DB 67033 ACATGTAAGACACACAAAGATCCAGATGCGCTAAAGATTTGTTATCTATTCAGCA 67092
 QY 1201 TGGCGACCGATTAACCCCTCCAGGTGCGCCCATATGAGATATTTGCCCGCAGAG 1260
 DB 67093 TGGCGACCGGTAATGATCTTATGTTGATGATGATGATGATGATGATGATGATGATGATGAT 67152
 QY 1261 GTACTACAGAAATTTGTTGCCAGGC---CCCGATGCTACTACACACCGATGTGGA 1317
 DB 67153 GCTATCAGGTTCTCTATCGTCTGACGAGGAGAAAGCTTATGTCACTTCAGATGTGG 67212
 QY 1318 CAACACCAATGTTGGGCGCGCCAGTTTGAAGATG---CCCGCGCATGAGTTTC 1374
 DB 67213 CAGCATGATGATTTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 67272
 QY 1375 AGTCTGCTTGGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1433
 DB 67273 TCCGATGTTTGGGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 67332
 QY 1434 AGTGGGGAAGACCGGTCATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1493
 DB 67333 AAACCGAAGCGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 67392
 QY 1494 GAATGGAACCTTAGCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1553
 DB 67393 GAGGTGTCACCGCATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 67452
 QY 1554 TGGCAGGAGTGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1613
 DB 67453 TTCTTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 67512
 QY 1614 AACATGCCAGGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1673
 DB 67513 TATATG---GAATCTTGTGATTTTGTGCAAAATTTGCTGAACTTACGATGATGATGATGAT 67569
 QY 1674 ATTAATGTCGCAAGCGGGAAGTTTGGCCCGGCGATGCGGAA 1718
 DB 67570 ATTTTATTTCAACGTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 67614

RESULT 6
 ACF65387_0
 WP Sequence split into 7 fragments LOCUS ACF65387 Accession ACF65387
 WP Fragment Name Begin End
 WP ACF65387_0 1 110000
 WP ACF65387_1 100001 210000
 WP ACF65387_2 200001 310000
 WP ACF65387_3 300001 410000
 WP ACF65387_4 400001 510000
 WP ACF65387_5 500001 610000
 WP ACF65387_6 600001 696798
 ID ACF65387 standard; DNA; 696798 BP.
 AC ACF65387;
 AC XX
 DT 20-NOV-2003 (first entry)
 XX XX
 DE Photobadus luminescens nucleotide sequence #40.
 XX XX
 KW Antibacterial; fungicide; insecticide; polymorphism; genetic analysis;
 KM detection; food; gene expression; plant; animal; microorganism; toxin;

QY 1494 GAACTGGGAACCTTACCCGATGACATCCAGTTAAACTATTTCTCAATACCGT 1553
DB 6906 GAGCTGTCCACCGCATTTGAAATATGGTTGCCGTTCTGTGAACTTGAATATCGT 6965
QY 1554 TGGCAGGGATGGTGGGTGAGTGGCAACAATTCTTACGAAAGAACGTTATCTGCTCT 1613
DB 6966 TTTCTGGGATGATTAAGACATGCGAAGATATGATTTATGCTGCTGCTGCTCAATCT 7025
QY 1614 AACATGTCCAGGGCATGCGCAATTAATCTCTCTGTAAGCCCTATGCGATCAAGGT 1673
DB 7026 TATATG---GAATCTTGGCTGATTTTGTCAATTTGGCTGAAGCTTACGATGATGGC 7082
QY 1674 ATTACGTGGCGAAGCGGAAAGTTTGGCCCGCGCATGCGCGAA 1718
DB 7083 ATTCTATTCACGTTATATGATGAAAGTAAATTTGCGCGAA 7127

RESULT 7
ACA39863

ID ACA39863 standard; DNA; 1878 BP.

AC ACA39863;

DT 19-JUN-2003 (first entry)

DE Prokaryotic essential gene #21520.

KM Antisense; ds; prokaryotic essential gene; cell proliferation;

XX drug design; gene.

OS Mycobacterium leprae.

PN WO20027183-A2.

PD 03-OCT-2002.

PF 21-MAR-2002; 2002WO-US009107.

PR 21-MAR-2001; 2001US-00815242.

PR 06-SEP-2001; 2001US-00948993.

PR 25-OCT-2001; 2001US-0342823P.

PR 08-FEB-2002; 2002US-00072851.

PR 06-MAR-2002; 2002US-0362699P.

PA (ELIT-) ELITRA PHARM INC.

PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;

PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

DR WPI; 2003-029926/02.

DR P-PSDB; ABUS5993.

XX New antisense nucleic acids, useful for identifying proteins or screening

XX PT for homologous nucleic acids required for cellular proliferation to

XX PT isolate candidate molecules for rational drug discovery programs.

XX Claim 14; SEQ ID NO 27733; 1766bp; English.

CC The invention relates to an isolated nucleic acid comprising any one of
CC the 623 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product, or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies

CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target
CC prokaryotic essential genes. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 1878 BP; 375 A; 548 C; 596 G; 359 T; 0 U; 0 Other;

Query Match 19.5%; Score 381.6; DB 7; Length 1878;
Best Local Similarity 54.3%; Pred. No. 4; de-110;
Matches 941; Conservative 0; Mismatches 759; Indels 34; Gaps 7;

QY 109 CTGAAACGTGTTCTTACGCAACCGGAGGCTTATCTCCGATGATGATAGCTGAAA 168
DB 109 CTGAAACGTGTTCTTACGCAACCGGAGGCTTATCTCCGATGATGATAGCTGAAA 168
QY 169 CGCCATGAGGATCAAAACATTTTGGCTATCCGCGCGGCAATTTTGGCCATCTATGAT 228
DB 169 GAACTCATGTCAGGATCTTCCGGAATCCCGCGGCTCCGCTCTGATATATGAC 228
QY 229 GAACTCATGTCAGGATCTTCCGGAATCCCGCGGCTCCGCTCTGATATATGAC 228
DB 229 GAACTCATGTCAGGATCTTCCGGAATCCCGCGGCTCCGCTCTGATATATGAC 228
QY 229 CGGCTGTTTACTGAAAA-----AAGTTAGCTATGCTGTCGCGCAAGACAG 279
DB 229 CGGCTGTTTACTGAAAA-----AAGTTAGCTATGCTGTCGCGCAAGACAG 279
QY 289 GAACTTCCCATGCGGCGGATGAGTATCCAGAGCCACAGTAAAGTGGAGTTGTTTC 348
DB 289 GAACTTCCCATGCGGCGGATGAGTATCCAGAGCCACAGTAAAGTGGAGTTGTTTC 348
QY 280 GGGCGCGGCAACCGGCGGCAAGGCTACGACAGTACCGGCAAGGAGGAGTGGATG 339
DB 280 GGGCGCGGCAACCGGCGGCAAGGCTACGACAGTACCGGCAAGGAGGAGTGGATG 339
QY 349 GGTTCATGTCAGCAGGCGGAGTAACTTGTGTAACCGGCAATTTGCGCAATTTGGAC 408
DB 349 GGTTCATGTCAGCAGGCGGAGTAACTTGTGTAACCGGCAATTTGCGCAATTTGGAC 408
QY 340 GCGACTTCTGCTCTGAGGCAACCAATCTGTGACACCGTTGGCGGATGCGATGAC 399
DB 340 GCGACTTCTGCTCTGAGGCAACCAATCTGTGACACCGTTGGCGGATGCGATGAC 399
QY 409 TCGGTGCGCAATGAGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 468
DB 409 TCGGTGCGCAATGAGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 468
QY 400 TCAATGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 459
DB 400 TCAATGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 459
QY 469 TTCCAGGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 528
DB 469 TTCCAGGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 528
QY 460 TTCCAGGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 519
DB 460 TTCCAGGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 519
QY 529 CGTATGCGGCGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 588
DB 529 CGTATGCGGCGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 588
QY 520 CGTATGCGGCGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 579
DB 520 CGTATGCGGCGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 579
QY 589 CGTATGCGGCGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 648
DB 589 CGTATGCGGCGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 648
QY 580 CGGCGGCGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 639
DB 580 CGGCGGCGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 639
QY 649 ATTCCCTCGACCCCGGATGATTAATTAATTAATTAATTAATTAATTAATTAATTA 708
DB 649 ATTCCCTCGACCCCGGATGATTAATTAATTAATTAATTAATTAATTAATTAATTA 708
QY 640 AG-----CTGGCACCGAAATGAGCTTCCCGGATGACAGCGCAACCAACCGCAC 693
DB 640 AG-----CTGGCACCGAAATGAGCTTCCCGGATGACAGCGCAACCAACCGCAC 693
QY 709 CCGGCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 768
DB 709 CCGGCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 768
QY 694 AACCGGAGATTCGTCGCGGAGCGCAAGCTGATGTCGCGCGCAAGCTGTCGTCGAT 753
DB 694 AACCGGAGATTCGTCGCGGAGCGCAAGCTGATGTCGCGCGCAAGCTGTCGTCGAT 753
QY 769 GTAGGGGAGGAGGAGTGGCGGCAATGCGGATGCGGCAAGTGGAGAAATTTGGGAAAG 828
DB 769 GTAGGGGAGGAGGAGTGGCGGCAATGCGGATGCGGCAAGTGGAGAAATTTGGGAAAG 828
QY 754 GTGCGGCGGCGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 813
DB 754 GTGCGGCGGCGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 813
QY 829 TTCCAGTGGCGGTAACCAACCGGATGAGGAAATTTGGGCTTTGACGAAACCAATGCC 888
DB 829 TTCCAGTGGCGGTAACCAACCGGATGAGGAAATTTGGGCTTTGACGAAACCAATGCC 888

Db 814 ACCGGCATCCGGTGGTTAGCAGTTGATGGCAGTGGCGGTTCCCGCAGACCATCAC 873
 QY CTTTCGGTGGTATGTTGGTATGCATGATGGCCACCGCTATGCGCACTTTGGCGTACGGAA 948
 Db 889 CTTTCGGTGGTATGTTGGTATGCATGATGGCCACCGCTATGCGCACTTTGGCGTACGGAA 948
 Db 874 CAGAACTTGGGATCCCGGCGATGACGGCAACGTTGGCGGGTGGAGCGCTGCACGCG 933
 QY 949 TGTGATTGTTGATTCAGTGGGGCCCGTTTCAGACACCGGGTAACTGGCAACTGAC 1008
 Db 934 AGCGATTATTCAGTGGCACTGGGCACTGCTTTTACGACCGGGTGAAGTGGCAACTGAC 933
 QY 1009 GAATTTGCTAGCCGGCCAAAGTAAATTCATTTGACATTCGACCCGGGCGAGTGGGAAA 1068
 Db 994 TCTTCGGCGCCGACCGCAAGCTTATACATTCGACATTCGACCCGGGCGAGTGGGAA 1053
 QY 1069 AACAGGCTCCCGATGTCCTCATTTGTGGGGGATGATGACCATGTTTAAACAGCTTTG 1128
 Db 1054 AACCGGACAGCTGACGTGCGCATTTGTGGGCGACCTCAAGCTGTTATGTAAGCTCATC 1113
 QY 1129 CAGCGGGCCCGGGAATGATTTACCCCAACCATCCCATACCCCGGAGATGGTTAAAT 1188
 Db 1114 GCGATCTGCGCATTAAGAGGTCCCGGCAACATCGAATGACGACTGGTGTCTAT 1173
 QY 1189 CGCATTTGATCATTTGGCGGACCGATTAACCCCTCGAGTGGCCCATATGAGATTAAT 1248
 Db 1174 TTGACGGGTGTCGCAAGACGATCCGTTGAGCTACAGTCCACAAAGCGACGACTGTG 1233
 QY 1249 GCGCCCGCAGAGTGTACAGAAATTTGT---CGCCAGCGCCCGCATCTACTACAC 1305
 Db 1234 AGCCGGAGTACGTATGAAAGCTTGTGAAATTTGTGGCCCGGAGCGGTATATGTC 1293
 QY 1306 ACCGATGTGGCACAACCAATGTGGCGGCGCAGT---TTTGAACAAATGGCGCCGC 1362
 Db 1294 GCGGCTGTAGTACGATCAGATGTGGCGGCGAGTTCATCTCTCGAAGAACCGCGT 1353
 QY 1363 CGATGATTTCCAGTGTGGCTTGGGTACATGGGCTTTGTTTAACTGCGCGCATAGGA 1422
 Db 1354 ACTTGCTCAACTCAGCGGTGTGGGACCAATGGGGTTGTATTCGCGCGCATAGGC 1413
 QY 1423 GCCAA-AGTGGAGTGGGGGACGAGCGGTCAATTTGATGAGAGATGCCAGTTCCAA 1481
 Db 1414 GCCAAGATTGCTCCCGCGGAGGACCAAGTCTGGCGATGACGCGATGCTGCTTCAG 1473
 QY 1482 ATGAATCTTCAGGAATCTGGGAACTCTAGCCCAATGACATCCAGTTTAAATCTATTAT 1541
 Db 1474 ATGACCAACCAAGAGTGGTACGTCGCTATCCAGGGCGCGCATCAAGTGGCGCTTG 1533
 QY 1542 CTCATTAAGGTGGCAGGGGATGGTGGTCAAGTGGCAAGAACTTTCTACGAAGACGT 1601
 Db 1534 ATCAACAAAGGCACTTGAGGCAATGGGCAATGGCAAGCTTTTATCAAGACGT 1593
 QY 1602 TATTCGCTTCAATG---TCCAGGGCATGCGACATTAATCTCTGT 1652
 Db 1594 TACTCGACAGACGACTGGCCACCACTCACACCGCATTCGGAATTGGAAGTTGGCC 1653
 QY 1653 GAAAGCTATGATCAAGGATTAATCT---GTGGCGAAGCGGGAAGATTGGCCCGCG 1709
 Db 1654 GAGGGGTGGTGGTGGTGGATGCGTGGCAAGTGTGAAGAGAGCTTGTGATGTATC 1713
 QY 1710 ATCGCGGAATGCTAGCCCAACATGATGCTGTGATGATGATGATGATGATGATGAT 1769
 Db 1714 AATCAGCGCAGGCGCATTAACAACCGCCCTGTGGTGAATGCACTTCATGTCGGCGCAT 1773
 QY 1770 GAAATCTGTACCTATGATTTGCCCCCGCATGAGTAATGCCCAAATGCTAAGT 1823
 Db 1774 GCGAGGTGTGGCCGATGCTGCGCCGCTGGGGCGAGCAAGCAGATCCAGACT 1827

RESULT 8
 ACA37690
 ID ACA37690 standard; DNA, 1866 BP.
 AC ACA37690;
 XX

DT 19-UTN-2003 (first entry)
 XX Prokaryotic essential gene #19347.
 DE Antisense; ds; prokaryotic essential gene; cell proliferation;
 KM drug design; gene.
 XX Mycobacterium avium.
 OS
 PN W020027183-A2.
 PD
 XX 03-OCT-2002.
 PF 21-MAR-2002; 2002MO-US009107.
 PR 21-MAR-2001; 2001US-00815242.
 PR 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.
 XX (ELIT-) ELITRA PHARM INC.
 PA Wang J, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW,
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH,
 PI
 XX
 DR WPI: 2003-029926/02.
 DR P-PSDB; AEU33820.
 XX
 PT New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 PS Claim 14; SEQ ID NO 25560; 1766bp; English.
 CC The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-regulated gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence data for this patent did
 CC not form part of the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 1866 BP; 313 A; 673 C; 624 G; 256 T; 0 U; 0 Other;
 Query Match 19.4%; Score 379.8; DB 7; Length 1866;
 Best Local Similarity 54.3%; Pred. No. 18e-109;
 Matches 938; Conservative 0; Mismatches 757; Indels 34; Gaps 7;

QY 109 CTGAAAACCTGTTCTAGCCCAACGCGGAGTATATCTGATGAGTATAGCTGAAA 168
 Db 100 CCGAAACGCAATTGGGCGCCGACAAAGTATATGATGCAATAGCTGATCCGATCGTGGAG 159
 QY 169 GCGCATTGGGGTCAACACATTTTGGCTATCCCGCGGGGCAATTTTGGCCATCTATATAT 228
 Db 160 GAGCTCGGGCTGAGAGTATCTTCGCAATTCGCGGCGGGCTCTCCGCTCTATATAC 219
 QY 229 GAATGTATCCGCTTTGAAACGCGGCGGAAATTTGACATATTTTGGTGGCCATGAACA 288
 Db 220 CCGCTGTGACT-----CAAAGACTGGGCAAGTCTGTGTCCGCCACAGACAG 270
 QY 289 GGAAGTTCCTCATGCGCGGATGGGTATGCGCAGAGCCCAAGTAAAGTGGAGTTTGTTC 348
 Db 271 GGTGCGCGATCAGCGCGCCAGCGCTACCGGCAAGCCACCGGCAAGGTGCTGTGCATG 330
 QY 349 GGTATCATCTGACCAAGGCGGCACTAATCTGTGACCGGCAATGGCAATGCCCATTTGGAC 408
 Db 331 GCCACTCCGGTCCGGGCGCCACCAACTGTGTACCGGCTGGCCGACGCCAAGTGAAC 390
 QY 409 TCGGTGCCATGATGATGATTACTGAGAAGGTGGCCGCTGCCATGATTTGATGCGATCT 468
 Db 391 TCGATCTCCGCTGTCGCGCTCACCGGCGAGGTGGAGCAGCGCTGATCGGCAACCGAC 450
 QY 469 TTCCAGGAATTTGACATTTTGGCATCCCTTACCGATCGTTAAGCACTCCATGTGTGTA 528
 Db 451 TTCCAGAGAGCCGCAATCTCCGCAATCCCATGCTGATCACCAACCAACTCTCTGTGTC 510
 QY 529 CGTATGCGCGGATGATGCTGCAATTTTACTGAGGCTTTCATCTTGTAGCACCGAT 588
 Db 511 CGCTCCGGACAGAAATCCCGAGGTGCTTGCAGAGGTTTCCATGCGCGCTCCGCGC 570
 QY 589 CGTCCCGGCGGCTTTTGTATGATATTTCCCAAGATGTTGGGCTTGAAGAAATGTGATGAC 648
 Db 571 CGCCCGGCGGCTCTGTCGACATCCCAAGAGCGTGCAGAGGCGCAGTGCAGCTTC 630
 QY 649 ATTCCCTCGACCCCGGATGATCTATCTAACCGGCTTATCGCCCGCAGGTTAAAGGTAT 708
 Db 631 AG-----CTGGCCCGCGGATCCACTGTCGCGGATCAAGCCGACCAACCAACCGAC 684
 QY 709 CCGGCAAAATTAATGCGGCAATTTGATGAGAGGCGGCAAAATCCCTGTCTTAC 768
 Db 685 AGCCCGAGATCCGGAAGCCGCCAAGCTGATCGCGCGCGCAACCGATGCTGTAC 744
 QY 769 GTAGGGGGAGGGGCGATGCGGCGCAATGCGCGAGGTCAGAGATTTTGGGAAAG 828
 Db 745 GTAGGCGGCGTGTATCTCGCGGAGGCGCAGCGAGCTGCGGAATCGCGAGCTG 804
 QY 829 TTCCAGTTGCGGTAACAACACCTGATGGAATTTGGGCTTTTGAAGAAACCATCC 888
 Db 805 ACCGCGATCCCGGTGATGACACGCTGATGCGCGCGCGGCTTCCGGAACGCAACGCG 864
 QY 889 CTTTGGGTGGTATGTTGGGTATGATGCGCACCGCTATGCGCACTTTGCCGTACGCCAA 948
 Db 865 CAGCACTTGGGATCCCGGATGCAAGCGACCGTGGCGGCGGCGCTTGAAGCGC 924
 QY 949 TGTGATTTGTTGATGATGAGTGGGGCGCGTTTTCAGACCGGTTACTTGGCAAACTAGAC 1008
 Db 925 AGCGACTGTGTATGCGCTGTGATGCGGTTTCAATGACCGGTGACCGGCAAGCTGAC 984
 QY 1009 GAATTTGCTAGCGCGGCGCAAGTAATTAATGATGACATGACCGCGCGAGGTGGAAAA 1068
 Db 985 ACCTTGCGCCGGAAGCCAGGTCAATCCACGCGCATGACCGCGCGAATCGGCAAG 1044
 QY 1069 AACAGGGCTCCGATGTCCTTGTGGGGGATGTAAGCCATGTTTGAACAGCTTTG 1128
 Db 1045 AACCGGATGCGGATGCTGCGGCACTGTGCGGCAAGCGCGCTATCGCGAACTGTTC 1104
 QY 1129 CAGCGGCGCGGAAATGATTAATCCCAACCTCCCATACCAACCGAGCATGTTAAAT 1168
 Db 1105 GAGATGCTGCGCAGAGCGGGCGCCGGAATCTGACATGCGGATGCTGTGGGCTTAC 1164
 QY 1189 CGCATGATCAATTGGCGAGCGAATTAACCCCTCCAGGTGCCCACTATGAGAGTATAT 1248

Db 1165 CTGACAGAGTCCAGTCCACTTATCCGCTAGCTACAGGGCGGAGAGCGGACGCGTC 1224
 QY 1249 GCCCCCCAGAGATGATGACAGAAATTTGTGCGCAGGCC--CCCATGCTCTATACAC 1305
 Db 1225 GGGCCGATGATCGTATCGAAGCTGGAGCAATGCGCGGCGGACCGCTGTACGTG 1284
 QY 1306 ACCGATGCGGACAAACCAATGTTGGCGGCGCAGT---TTTGAACAATGCGCCCGC 1362
 Db 1285 GCGGCGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1344
 QY 1363 CGATGATTTCCAGTGTGCTTGGGTGATGATGATGATGATGATGATGATGATGATGAT 1422
 Db 1345 ACCGTGCTCAATTCGCGCGCGCTGGGCAACCATGAGGTTCCCATTCGCGCGGATGGGC 1404
 QY 1423 GCCAAGTGGG-AGTGGGGAGAGAGGCGTCAATTTGCATGATGAGATGCGACCTTCCAA 1481
 Db 1405 GCCAAGTGGGCGGTCCGAGGCGCGAGGTGCGGCGATGCGGCGAGCGGTTGTTCCAG 1464
 QY 1482 ATGAACTTTCAGGAATGGAACCTTACCCAGTACAGATCCAGGTTAAACTATATAT 1541
 Db 1465 ATGACCAACGAGAGTGGCCACTGGCGCGTGGAGGGGATCCGATCAAGGTGGCGCTG 1524
 QY 1542 CTCAATPACGTTGGGAGGAGTGGCGTCACTTGGCAACAACTTTTACGAAGAAGT 1601
 Db 1525 ATCAACACGGAACCTGGGCAATGTCGCGCAATGAGACCTGTGTTACGAAGAGCGG 1584
 QY 1602 TATTC-----TGTCTTAAATGTCGCCAGGAGATGCGACATTAATCTCTCTG 1652
 Db 1585 TATTCGAAACGAGATGCGCACCACTGCGACCGGATCCGATCTTGTCAAGCTGCGC 1644
 QY 1653 GAAGCTTATGCGATCAAGGCTAT--TACTGTGCGAAGCGGGAAGTATTTGGCCCGGCG 1709
 Db 1645 GAGCGGTGGGCTGCGTGGATGCTGCGATGCGAGCTGAGGAAACCTGTGACGATGATC 1704
 QY 1710 ATGCGCGAATTTGTAAGCCCAATGATGCTGTGATGATGATGATGATGATGATGATGAT 1769
 Db 1705 AAGCTGCGCGGCGATCAACGACCGGCGGCTGTGATGATGATGATGATGATGATGATGAT 1764
 QY 1770 GAAACTGTATCCCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1818
 Db 1765 GCGCAGGTGTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1813

RESULT 9
 ID ACA53435 standard; DNA; 1722 BP.
 ACA53435;
 AC ACA53435;
 XX AC
 DT 19-JUN-2003 (first entry)
 XX DT
 DE Prokaryotic essential gene #35092.
 XX DE
 KM Antisense; ds; prokaryotic essential gene; cell proliferation;
 XX KM
 OS Vibrio cholerae.
 XX OS
 EN W0200277183-A2.
 XX EN
 PD 03-OCT-2002.
 XX PD
 PF 21-MAR-2002; 2002MO-US009107.
 XX PF
 PR 21-MAR-2001; 2001US-00815242.
 XX PR
 PR 06-SEP-2001; 2001US-00948993.
 XX PR
 PR 25-OCT-2001; 2001US-0342923P.
 XX PR
 PR 08-FEB-2002; 2002US-00072851.
 XX PR
 PR 06-MAR-2002; 2002US-0362699P.
 XX PR
 PA (ELIT-) ELITRA PHARM INC.
 XX PA

PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Foreyth RA, Xu HH;
XX MPI; 2003-029926/02.
DR P-PEDB; ABU49565.

PT New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.

PS Claim 14; SEQ ID NO 41305; 1766bp; English.

CC The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target
CC prokaryotic essential genes. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pcc_sequences

XX Sequence 1722 BP; 471 A; 374 C; 451 G; 426 T; 0 U; 0 Other;

XX Query Match 19.4%; Score 379.6; DB 7; Length 1722;
XX Best Local Similarity 53.9%; Pred. No. 2e-109;
XX Matches 917; Conservative 0; Mismatches 764; Indels 19; Gaps 6;

QY 128 AACGGCAAAACGGGGCTTATCTCTGATGATACCTGAAAAGCCATGGGGTCAACACA 187
DB 5 AAATCTATCTGGCGCAGAGATGATCGCAATCTGATCAATGAAGGTGAGCAAA 64
QY 188 TTTTGGCTATCCCGGGGGCAATTGGCCATCTATGATGAAGTGAACGGCTTGAAG 247
DB 65 TCTTGGATATCCCGGGGGTCAAGTCACTGACATCTGAGTCCCTTCAAC-----GAAA 118
QY 248 CGGCGGGGAAATTGAGCATATTTTGTGCGCCATGAACAAGAGCTTCCCATCGGCGG 307
DB 119 AAAGAGACCAATTAAACAGCTTGGTTCGCAAGAAAGCGCAACATATGCTG 178
QY 308 ATGGGTATCCAGAGCCACAGTAAAGTGGAGTTTGTGGTACATCTGAGAACAGGG 367
DB 179 ATGGCTATGCGGCGCCACTGAAAAACGGGGCGTAGTCTGTCTGCTCAAGTCCGGGG 238
QY 368 CGACTAATCTGGTGACCGGCATTTGCAATGCGCATTTGGACTCGGATCGGATGGTGA 427
DB 239 CCACCAATACGCTGATGAGATGGAACCGCTATATGAGATCGATTCAGATGATTTGA 298
QY 428 TTACTGGAAGGTGGGCGGTGCATGATTGGTAGCATCTTTCCAGAAATTTACATTT 487
DB 299 TCTCGGTACGTAGCCCACTAATCTTATTTGTAAGATGACATTTCAAGAGTGCATATCG 358

QY 488 TTGGCATACCTTACCGATCGTTAAGCACTCCATGTGTGTAAGTATGCGGCGATATGCG 547
DB 359 TCGGGGTGTCTGTCCCATCGTTAAACAGGTTTTTGGTAAAAAAGCCAGAAATATTC 418
QY 548 CTCGATTTGTAAGAGGCTTCCATCTGTGACGACCGGTCTGCCGGCGGTTTTGA 607
DB 419 CTGAAGCATCAAAAAAGCGTTTTATATGCGTCAACGGGGGCTCTGTCTGTAGTGA 478
QY 608 TCGATATTTCCAAAGATGTGGGCTTGAAGAAATGTAGTACATTCCTCCAGCCCGGTG 667
DB 479 TTGACGTGCAAAAATGATGATGAGCCCGCTTAAATACCTGCTTATGATCCAGAAA 538
QY 668 AGCTTATCTACCGGGTTATGCGCCAGGTTAAAGATATCCCGCAAAATTAATGCGG 727
DB 539 CATTAAATGCGCTTTTAAACCAACAGACGCTGCGCAAAAGAGATCAAAAAG 598
QY 728 CATTGCAATTTTGGAGACAGCCAGAAATCCCTTGTCTAGCTAGGGGAGGCGGATCG 787
DB 599 GGTACGGCACTGCTCGATGCGAAAAAGCCAGTACTAGTGTGCGTGGCGGGCGATTA 658
QY 788 CCGCAATGCCATGCCAGGTCAGAAATTTGCGGAAAGGTCCAGTTGCGGTTACAA 847
DB 659 TTGCTAATGACATGACGAAATTCGCAACTTGCAGAGGCTTAAATCTGCCCGTTA 718
QY 848 CCACCCGTATGGAATTTGGGCTTTTGAAGAAACATCCCTTGGGTGGATATGTTG 907
DB 719 GTACTTTATGATGATTTGGGCGTTTCTGGGACTATAAAAACCTCACTGGGTATGCTG 778
QY 908 GTATGATGCGCAACCGCTATGCGAATTTGCGGTAGCGAATGTGATTTGTGATGCA 967
DB 779 GCATCATGCTGTGTATGATGAAGCAATATGGGATGCAAAATGCAATCTGATTTTGGG 838
QY 966 TGGGGGCGCGTTTGCAGACCGGGTAACTGAGCAAAATTTCTAGCCGCGCA 1027
DB 839 TAGGTGTGCGCTTTATATACGCTAGCATTAACAACCTGAGAAAGTACCCCAAGCGA 898
QY 1028 AAGTATTCATTAATGACATGACCGCGGAGTGGGAAAAAAGCGCTCCGATGTGC 1087
DB 899 AAATCATGATATCATATTTGACCGGTATGATTTCTAAAAAGCTGAATGATCTGCG 958
QY 1088 CCATGTGGGGGATATGAGCCATGTTTGAAGACTTTTCAGCGGGCGCGGGAATGG 1147
DB 959 CGATTTGTGTCTGCGCAAGATTTGATGATGATGATCTCAAGCTTTGGAAGAGAGCG 1018
QY 1148 ATTACCCCAACCCATCCCATACCAACAGGATGTTAAATGCAATGATCATTTGGCGGA 1207
DB 1019 CTGAGGCAATATGCGGAGCATTTGATGATGATGATGATGATGATGATGATGATGATG 1078
QY 1208 CCGATTTACCCCTCCAGGTGCGCCCATATGAGATATCTATTTGCCCCAGGA---GCTAG 1264
DB 1079 ATGCTAGTGTGTGCTTATGAACTCAAGCAAGAGATTTAAGCTTACGCAAGTATG 1138
QY 1265 TACAGCAATTTGTGCGAGGCGCCGATGCTTACACCAACGATGTGGAGCAACACC 1324
DB 1139 AAGCGTATATTAATCAAGAGAGGTAAAGCAATTTTGGCTCGGATGTGGGGGAGCAAC 1198
QY 1325 AAATGTGGCGGC---CGAGTTTGAACATGCGCCCGCGGATGATTTCCAGTGTG 1381
DB 1199 AAATGTGTGCGGCACTTACTATCTGTTTGAATGAGCCGCTGAGGATTAATCTGTGCG 1258
QY 1382 GCTTGGGTACGATGAGGCTTTGTTTACCTGCGGCACTGAGAGCCAAATGAGGAGTGGGG 1441
DB 1259 GGTAGGCACTATGAGTTTGTGTTTACTGCGGGAATGGCGGTGAATTTCCGCAAGCCAG 1318
QY 1442 ACG-AAGCGTATTTGATCAATGAGATGCAAGCTTCCAAATGATCTTACAGAACTGG 1500
DB 1319 AAGAAGAGGTACTGTGCTGAGACCGAGATGAGATTCAGATGATATTTCAAGAGCTAT 1378
QY 1501 GAACCTAGCCAGTACGACATCCAGGTAAACATATATCTCATATACGATGAGTGGCAGG 1560
DB 1379 CGACGCACTGAGTACGACATTCAGTAAATATATTAATCTCAACAAACGTTTCTTG 1438

Db 823 GCGTCGCTGAACTGAAACCAAGTTCGGGACCTGCTCGGCTACCCGGTCAAGAACACGCTG 882
 QY 856 ATGGGAATTTGGGCTTTTGAAGAAACCATCCCTTCGATGGGATGTTGGTATGCAT 915
 Db 883 ATGGGCTTCGGGGTATCGCGCTGCGAAGAAATTCCTCGGCATGCTCGGCATGCAC 942
 QY 916 GGGCCCGCTATGCCAATTTGGCGTCAGGCAATGTATTTGTGATTTGCAATGGGGGCC 975
 Db 943 GGCACCTACGAAGGCAATGCGATGCAAGCATGCGATCGATCGATCGATCGATCGATCG 1002
 QY 976 GCTTCGACGACCGGGTAACTGGGCAAACTAGACGAATTTGCTAGCCGCGC---CAAGTA 1032
 Db 1003 CGCTTCGACGACCGGGTATCGCGGACCCGCGGACCTTCGCTCGCGCCCGCGCAATGC 1062
 QY 1033 ATTCAATTTGATGATGACCGCGGAGTGGGAAAAACAGGGCTCCCGATGTCCTCAT 1092
 Db 1063 ATTCAATTTGATGATGACCGCGGAGTGGGAAAAACAGGGCTCCCGATGTCCTCAT 1122
 QY 1093 GTGGGGGATGTACGCGATGTTTGTAGAACAGCTTTTGGAGCGGGCGCGGAAATTTGAT 1152
 Db 1123 GTGGGGGATGTACGCGATGTTTGTAGAACAGCTTTTGGAGCGGGCGCGGAAATTTGAT 1182
 QY 1153 CCCACCCATCCCATACCAACCGGATGTTAAATGCGATTTGATGATGATGATGATGAT 1212
 Db 1183 GGGCCCGACACCGAAGCGCTCGCCCAATGTGGAAGGACATCGAAGGCTGGCGCGGAG 1242
 QY 1213 TACCCCTCCAGGTCGCCCATATGAGGATCTATTCGCCCCCGGAGGTAGTACAGGAA 1272
 Db 1243 GACTGCTGAAATGACGACCGGAAACGATATCATACCGCGATGTCGATGTCGATG 1302
 QY 1273 ATTGTCGCGACGCGCGCGCG---ATGCTTACTACACACGATGTTGGGACAAACCAAT 1329
 Db 1303 GCGTGGAGCTGACGGAAGGCAATGCGTTCGTCGTCGACGTCGCGCGACACCAATG 1362
 QY 1330 TGGCGCGCGCGGTTT---TGAACATGCCCCCGCGGATGTTTCAATGCTGCGTTG 1386
 Db 1363 TGGCGCGCGCGGTTT---TGAACATGCCCCCGCGGATGTTTCAATGCTGCGTTG 1422
 QY 1387 GGTACGATGCGCTTTGTTTACTGCGCCCATGAGGACCAAGTGGAGTGGGAGTGGGAG 1445
 Db 1423 GGCACGATGCGCTTCGCGCTGCGCGCGCGCATGGGCGTCAAGATGCGCCATCGGAGAC 1482
 QY 1446 GCGGCTATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1505
 Db 1483 GACGCTGCTGATCATCGCGGAGGCTGATTCAGATGTCATCCAGGAATGTCGATC 1542
 QY 1506 CTAGCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1565
 Db 1543 TGCTGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1602
 QY 1566 GTGCGTCAAGGCAAACTTTTACGAGAACGTTATTCGTTTAAATGTCGATGTCGATG 1625
 Db 1603 GTCCGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1659
 QY 1626 GGCATGCGACATTAATCTCTCTGAGAGCCATGAGCATGAGGATATTAATGTCGATG 1685
 Db 1660 GCGCTGCGATTTCTGTAAGCTGCGGAGCGATGCGGATGCGGATGCGGATGAA 1719
 QY 1686 AAGCGGGAAGTTTGGCGCGCGATGCGCGGAAATGCTAGCCCAATG---GTCTGTG 1742
 Db 1720 AAGACTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1779
 QY 1743 GTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1802
 Db 1780 TTTCTGATTTCCAGACCGATGCGAGGAAACGTCGCGGATGATGATGATGATGATG 1839
 QY 1803 AGTAATGCGAAATGCTAGTTTACCGGAAATGCGG 1838
 Db 1840 GGCATCAACGAGATGCTGCTCGATCGGAATGATCTG 1875

RESULT 11
 AAS54310

ID AAS54310 standard; DNA; 1725 BP.
 AC AAS54310;
 XX
 DT 13-FEB-2002 (first entry)
 XX
 DE Pseudomonas aeruginosa DNA for cellular proliferation protein #441.
 XX
 KW Antisense; ds; prokaryotic cellular proliferation gene; antibiotic;
 KM antibacterial; drug design.
 OS Pseudomonas aeruginosa.
 XX
 PN MO200170955-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 21-MAR-2001; 2001WO-US009180.
 XX
 PR 21-MAR-2000; 2000US-0191078P.
 PR 23-MAY-2000; 2000US-0206848P.
 PR 26-MAY-2000; 2000US-0207727P.
 PR 23-OCT-2000; 2000US-0242578P.
 PR 27-NOV-2000; 2000US-0253625P.
 PR 22-DEC-2000; 2000US-0257931P.
 PR 16-FEB-2001; 2001US-0269308P.
 XX
 PA (ELIT-) ELITRA PHARM INC.
 XX
 PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
 PI Yamamoto RT, Xu HH;
 XX
 DR WPI; 2001-611495/70.
 XX
 PT P-PSDB; AAN36451.
 PT
 PT New polynucleotides for the identification and development of
 PT antibiotics, comprise sequences of antisense nucleic acids.
 XX
 PS Claim 27; SEQ ID NO 7947; 511p; English.
 XX
 CC The invention relates to antisense inhibitors of genes essential to
 CC prokaryotic cellular proliferation, their use in identifying the genes,
 CC their use in the discovery of novel antibiotics, the essential genes,
 CC themselves and the encoded proteins. The prokaryotes used are Escherichia
 CC coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,
 CC Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
 CC useful for the identification of potential new targets for antibiotic
 CC development. The antisense nucleic acids can also be used to identify
 CC proteins used in proliferation, to express these proteins, and to obtain
 CC antibodies capable of binding to the expressed proteins. The proteins can
 CC be used to screen compounds in rational drug discovery programmes. The
 CC antisense nucleic acid sequence is also useful to screen for homologous
 CC nucleic acids which are required for cell proliferation in a wide variety
 CC of organisms. The present sequence encodes an essential prokaryotic
 CC cellular proliferation protein. Note: The sequence data for this patent
 CC did not form part of the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pat_sequences
 CC
 SQ Sequence 1725 BP; 351 A; 546 C; 517 G; 311 T; 0 U; 0 Other;
 XX
 Query Match 19.1%; Score 373.8; DB 4; Length 1725;
 Best Local Similarity 54.6%; Pred. No. 1.4e-107;
 Matches 866; Conservative 0; Mismatches 712; Indels 25; Gaps 6;
 QY 135 AACCGGGCTTATATCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 194
 Db 12 ATCTGCGCGTGAATGTCGTCGCTCGTTCGCGGACGACGAGCGTTAGTACATCAACG 71
 QY 135 CTATCCCGCGCGGCAATTTTGGCCATGATGATGATGATGATGATGATGATGATGATG 254
 Db 72 GTACCGGAGTGTGCTCTCTGATATCTACGACGCGCTTTC-----AAGAGCA 122

QY 255 GGAATTGAGCATATTTGGTGGCCATGAAACAAGAGCTTCCATGGGGCGAGTGGTA 314
 Db 123 GCACTGACCCACATCTGTGGTCCGACGACAGCGGGCTACCTTAATGGCCGAGGCTA 182
 QY 315 TGCCAGAGCCAGAGGTAAAGTGGAGTTTGTTCGGTACATCTGAGACGAGGCGACTAA 374
 Db 183 GCGCCGCGCCACCGGCAAGCCCGGTGTGTGTGTGACTTCGCGCCGGGGGGGACCA 242
 QY 375 CTGTGTGACCGGCAATTTGCCAATGCCATTTTGGACTTCGGTGGCCATGGTGTGATTACTGG 434
 Db 243 GCGCATCACCGGTATCGCACCGCATACATGATGATCCCGATGGTGTGATCTGTTCGG 302
 QY 445 AGAGGTGGCCGCTGCGCATGATGTGATGAGATGCTTTCAGAGAAATTTGATTTGGCAT 494
 Db 303 CCAAGGTGCCAGCAACATGTGTGATCCGATCCGTTCCAGAGAAACCGACATGTGCGGAT 362
 QY 495 CACCTTACCGATGTGTTAAGCACTCTTATGTGTACGTATGCGGCGGATATGTCTCGCAT 554
 Db 363 CTCCCGCGCGATCGTAAAGCACAGCTCATCATCAAGCATCCCTGGAAATTTCCGAGGT 422
 QY 555 TGTATCTGAGGCTTTCATCTTGTAGACACCGGTGTCCCGGCGCGTGTGATGATAT 614
 Db 423 GATCAAGAGAGCTTCTTACCTGTGGCGAGTCCGATGTCCGGGCGGCTGTGTGTGATAT 482
 QY 615 TCCCAAGAGTGTGGCTTAAAGAAATGTAGATACATTCCTCCGTCGACCCCGGTGACGTTAA 674
 Db 483 TCCGAGAGCATGGGCGAGCCGACGCAAGAAATTCGATATTCCTATCCGAGAAAGGTCA 542
 QY 675 TCTACCGGCTTATCCGCCCCAGCGTTAAAGTATCCCGACAAATTAATGGGCGATTGCA 734
 Db 543 GTTGGCTGTGACACCGCGCGCTTCCGGGTCACTCGGAGACATCCGCAAGGCGCGCA 602
 QY 735 ATTGTGAGCAGCAGCAGAAATCCCTGTCTACGTAGAGGAGGCGGATGCGCGCA 794
 Db 603 GATGCTCTGGCGCGCAAGCGCCCGTGTGTATTCGCGCGCGCGCGTGTGTGTGAGCA 662
 QY 795 TCCCATGCGCAGGTGACAGAAATTTGCGAAAGTTTCCAGTTGCGGTTAACAACACCT 854
 Db 663 TGCCCGCGCGCTGACGAGCTGCGCAGATGCTCAACCTGCGCGGTGACCAACACCT 722
 QY 855 GATGGAAATTTGGGCTTTTGAAGAAACCATCCCTTCCGAGGATGTTGGGATGCA 914
 Db 723 GATGGGCTTCCGCGGATATCCGCGCGACACCGCATCTCTCGGCAATGCTCGGATCA 782
 QY 915 TGCGCACCGGTATGCAACTTTGCGGTGACGCAATGTATTTGATTTGATGAGTGGGCGC 974
 Db 783 TGCGCAGCTTCAACCGCAACTGCGATGATCAACGACGATGATCTCGCGTGGGCGC 842
 QY 975 CGGTTTCAAGCAGCGCGGTAACTGCGAAACTAGACAAATTTGCTAGCCCGCAAGTAA 1034
 Db 843 GGGTTTCAGCAGCGGATCATCAACGTGCGGCAAGTTCTGCGCAACGCGCAAGATAT 902
 QY 1035 TCACATTGACATGACCCCGCGAGAGTGGAAAAACAAGGCTCCCGATGTGCCATTTGT 1094
 Db 903 CACATTCGACATGATCCGGCGGTGATTTCCAAAGCATCAAGGCGCATCCGATGCT 962
 QY 1095 GGGGGATGACCGCATGTTTGAACAAGCTTTTGAAGCGGCGCGCGGAATTTGATTAACC 1154
 Db 963 CGGCGCGGTGACACCGTTTCAACCGAAATGTGCGCATGTCAGAGGAAATTCGGGAGAC 1022
 QY 1155 CACCATCCCATACACCCAGCATGTTAAATGCAATGATCTATTGGC---GGACCGA 1211
 Db 1023 CCGGAACAGAGATGCCAGGCTGCTGTGTGAAGAGATGACAGATGCGGTGTGTAACG 1082
 QY 1212 TTAACCCCTCCAGTGGCCCACTATGAGATATCTATTCGCCCCCGAGAGTATGACGA 1271
 Db 1083 TGGGTGTTCCTCGTACGACAGGCGAGCGGACATCATCAAGCGCGAGACCGGTATCATG 1142
 QY 1272 AATTGGTGGCGAGGCGCC-----CGATGCTTACACACCGATGAGGAGCAACACCA 1335
 Db 1143 GACGCTTACAGAGTTTACCATGAGGATGCTTCACTACCTTCGATGTGCGGCGACACCA 1202
 QY 1326 AATGTGGCGCGCCAGTTTGA---CAATGGCCCCCGCGATGATTTCCAGTGTGG 1382

Db 1203 GATGTTGGCTTCGAGTACTACAAAGTTCAACAGCCCATGCTGTGATCAATCCGCGG 1262
 QY 1383 CTGGGTAGCATGAGGCTTTGGTTTACCTGCGGCATGGAGCCAAAGTGGAGTGGGGA 1442
 Db 1263 TCTCGGACCATGAGGCTTGGCTTCCGCGCGGAGTGGGATCAAGCTCAACTTCCCGA 1322
 QY 1443 CGA-GCGGTATTTGATCATGTGAGATGCGACTTCCAAATGAATCTTCAAGAACTGG 1501
 Db 1323 CGAGATGTGCTGTGTGACCGGCGAGGCAAGTATCCAGTAAATTCAGAACTGTG 1382
 QY 1502 AACCTTACCCAGTATGACATCCAGTTAAACTATTTATCTCAATTAACGTTGGAGGG 1561
 Db 1383 TACCTGCTGTGACATGACACTGCGGTGAAGATCGTCAACTGAACATGTGCTGCGG 1442
 QY 1562 GATGTGCTGTGAGTGGACAAACCTTTACGAGAAAGTTATTTGCTTCTTAACATGTC 1621
 Db 1443 CATGTTCGCGCAGTGGCAGACATGACATCAACACCGCTTATTCGATTCCTACATG-- 1500
 QY 1622 CCAAGGATGCGAGCATTAATCTCTGTGTGAAGCTTATGACATCAAGGTTATCTGT 1681
 Db 1501 -GATTCCTTCCGAGCTTGTCTCAAGCTTGTCCGAAAGCTTACGAGCATGTGCGCAT 1559
 QY 1682 GCGCAAGCGGAAAGATTTGGCCCGCGGATGCGCGAAATGCTAGCCCAAGTGTCTGT 1741
 Db 1560 CACGACCTGAAGATCTGAAGCCGAAGATGAGAGAGGCTTGCATGAAGAAATGCGCT 1619
 QY 1742 GGT 1744
 Db 1620 GGT 1622

RESULT 12
 ACA42654
 ID ACA42654 standard; DNA; 1725 BP.
 XX ACA42654;
 AC 19-JUN-2003 (first entry)
 XX
 DT
 DE Prokaryotic essential gene #24311.
 XX
 KW Antisense; ds; prokaryotic essential gene; cell proliferation;
 KW drug design; gene.
 XX
 OS Pseudomonas aeruginosa.
 XX
 PN W0200277183-A2.
 XX
 XX 03-OCT-2002.
 PD
 XX
 PF 21-MAR-2002; 2002W0-US009107.
 XX
 PR 21-MAR-2001; 2001US-00815242.
 PR 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0342933P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.
 XX
 PA (ELIT-) ELITRA PHARM INC.
 XX
 PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zykkind JW;
 PI Wall D, Twaddick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 XX
 DR WPI; 2003-025926/02.
 DR P-PSDB; ABU38784.
 XX
 PT New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 XX
 PS Claim 14; SEQ ID NO 30524; 1766pp; English.

CC The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target
CC prokaryotic essential genes. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_poc_sequences

XX Sequence 1725 bp, 351 A; 546 C; 517 G; 311 T; 0 U; 0 Other;

Query Match 19.1%; Score 373.8; DB 7; Length 1725;

Best Local Similarity 54.6%; Pred. No. 1.4e-107;

Matches 886; Conservative 0; Mismatches 712; Indels 25; Gaps 6;

QY 135 AACCGGCGCTTATTCCTGATGATAGCCTGAAGAGCCATGAGGCTCAACATTTTGG 194
DB 12 ATTCGGCGCTTAATATGTCGTCGCTGTCGCGAGCAAGGCGCTTAAGACATCTACGG 71
QY 195 CTATCCGCGGCGGCAATTTTGGCCATCTATGATGAATCTGACCGCTTTAAGCGCGGG 254
DB 72 GTACCCGCGGTGTCCTCTCTGATATTTAGACGCGCTCTTC-----AAGAGCA 122
QY 255 GGAATTTAGCATATTTTGTGTCGCATGAAAGAGAGCTTCCGATCGCGGATGGGTA 314
DB 123 CGACGTGACCAATCCGTCGTCGCGCAAGCAAGCGGCTAACCATATGCGCGCACTA 182
QY 315 TGCAGAGCCACAGGTAAAGTGGAGTTTGTTCGGTACATCTGAGACAGAGGCGGATA 374
DB 183 CGCCGCGCGCACCGCAAGCCCGGTGTGTCGTGTATCTTCGCGCGCGGCGAGCAAA 242
QY 375 CTGTGTGACCGCGCATTTGCCATATGCGATTCGATCGGTCCTATGCTGATTAATCTG 434
DB 243 CGCATATCACCGGTATTCGCAACCGCATATGATCTCCATCCCATGTTGATTCGTGG 302
QY 435 AGAGGTGGCCGTCGTCATGATTTGGATGATGCTTTCCAGAAATTGACATTTTGGCAT 494
DB 303 CGAGTGGCGAAGCAACATGTCGATCGATGCGTTCCAGAAACCAACATGTCGCGAT 362
QY 495 CACCTTACCATGCTTAAAGCACTCCTATGTCGATAGTAGTGGCGGATATGCTGCAT 554
DB 363 CTCGCCGCGCATCTGAGAGCAACAGCTTCATATAGACATCCCTCGAATTTCCCGAGT 422
QY 555 TGTACTAGAGCTTTTCATCTTGTCTAGACACCGTGTCCGCGCGGCTTTTGTATGATAT 614
DB 423 GATCAAGAGAGCTTTCTACCTGCGCGAGTCGTCGTCCGCGGCGGTCGTGTCGTATAT 482
QY 615 TCCCAAGGATGTGGGCTTAGAAGATGAGTATTCCTCCGACCCCGGTCGTCGTTAA 674
DB 483 TCCGAGGACATGGCGCAGCCGACGAGAAATTCGAATTTCTATCCGAGAGAGTCAA 542

QY 675 TCTACCGGTTATCGCCCAAGCGTTAAAGTATCCCGACAAATTAATGGCGCATTCGA 734
DB 543 GTTGGGTTCTGACACCCCGCGCTTGCCTGCTACCTGGAGACATCCGCAAGCGCGCA 602
QY 735 ATTGTGAGCAGAGCCAGAAATTCCTTCTACCTAGAGGAGAGGCGATCGCCGCA 794
DB 603 GATGCTCTCGCGCGCAAGCGCGCGGTGTATTCGCGCGCGCGGTATCATGGGCA 662
QY 795 TCCCATGCGCAGGTGCAAGAAATTTGGGAAAAGTTCCAGTTTCGCGGTAAACACCT 854
DB 663 TCCCGCGCGCGCTGACCGAGCTGGCGCAAGTGTCACTGCGCGGTGCAACACCT 722
QY 855 GATGGAATTTGGGCTTTTGAAGAAACATCCCTTTCGCTGGGTATGTTGGTATGCA 914
DB 723 GATGGGCTCGCGCGGTATCCCGGCGAGACCGGACATTCCTCGGCATGCTCGGATCA 782
QY 915 TGGCAACCGCTATGCAACTTTCGCTGAGGATGTATTTGTTGATGCAAGTGGGCG 974
DB 783 TGGCAGCTTACCGCGCAACTGCGCATCATACAGCAAGCTGATTCCTCGCGGTGCGCG 842
QY 975 CGTTTTCAGCAGCGCGGTATCTAGCAAACTATGCAATTTTGTAGCCGCGCAAGTAT 1034
DB 843 GCGTTTCAACAGCGGCTCATACAGTGGCGGCAAGTTTTCGCGCAAGCCGACATCAT 902
QY 1035 TCACATTCATTCGACCCGCGGAGTGGGAAAAACAAGGCTCCGATGTGCCATTTGT 1094
DB 903 CCACATTCGACATTCGACCGGCGTGCATTTCCAGACATTCAGGCGCGCATCCGATGCT 962
QY 1095 GGGGATGACCGCATTTTAAAGACATTTTGAAGGCGGCGCGGAAATGATTAACC 1154
DB 963 CGGCGCGGTGAGACCGGCTTCTACCGAATGCTGCGATCTCATGGAATTCGCGCAAC 1022
QY 1155 CACCATTCCTCATACCCACCGCATGTTAAATGCAATTCATTCATTCGCGCGGCA 1211
DB 1023 CCGAAGCAGAGATGCGCGAGGCTGCTGTTGAGAGATTCAGCAGTGGCGGTGTAACG 1082
QY 1212 TTAACCCCTCCAGGTCCTCATATGAGATATCTTTCGCGCGGAGGATGATACAG 1271
DB 1083 TGGGTTGTCCTGACGACAGAGCGGAGGAGCATTCAGGCGCGACCGTCATGGA 1142
QY 1272 AATTGATGCGCAGGCGCC-----CGATGCTTACTACACCAAGTGGGACAAACCA 1325
DB 1143 GACGCTACAGAGTTATCCATGCGATGCTTATACCTTCGATGTCGCGCGACACCA 1202
QY 1326 AATGCGGCGCGCCAGTTTGA--CAATGCGCGCGCGGATGATTCATGATGCTGG 1382
DB 1203 GATGTCGCTGCGCAGTACTACAGTTCAACAGCCCATGCTGAGTCAACTCGCGCG 1262
QY 1383 CTGGGTAGATGAGGCTTGTGTTTACTGCGCGCCATGAGGACCAAGTGGAGTGGGGA 1442
DB 1263 TCTCGGACATGAGCTTCGCTTCCGCGCGGATGAGGATCAAGCTCAACTTCGCGGA 1322
QY 1443 CGA-GCGTCATTTGATCACTGATGAGATGCAAGCTTCCAAATGAATCTTCAGAACTG 1501
DB 1323 CGACATGTCGCTGCGTACCGGCGAGGAGGAGTATCCAGATGAATCCAGGAATGTC 1382
QY 1502 AACCTTACCCAGTACGATTCAGGTTAAACTATTTCTCAATAACGTTGGCGAGG 1561
DB 1383 TACCTGCTGCTGACGACTGCGGTGAAGTGTCACTGAACAATGAGTCCCTGGG 1442
QY 1562 GATGTGCTCAGTGGCAACAACTTTCTAGAGAAAGTATTTCTGCTTAAATGTC 1621
DB 1443 CATGTGCGCAAGTGGGAGGACATGCAATACAGCGGTATTCGATTCCTCATG- 1500
QY 1622 CAGGGCATGCGCAATTAATCTCTCTGGAAGCTTATGATCAAGGATATTAAGT 1681
DB 1501 -GAATCCCTGCGGACTTGTCAAGCTTGCAGAGCTTACGGAAGCTTACGAGTCCGAT 1559
QY 1682 GCGCAAGCGGAAAGTTGGCCCGCGCGGACGGAATGCTAGCCCAATGATCTGT 1741
DB 1560 CACGACCTGGAAGATCTGAAGCGGAATGAGAGGCGCTTCCGATGAAGATCGCT 1619

QY 1742 GGT 1744
 DB 1620 GGT 1622

RESULT 13
 AAT45893
 ID AAT45893 standard; DNA; 1734 BP.
 XX AAT45893;
 AC
 XX 12-FEB-1997 (first entry)
 DT
 XX
 XX R. capsulatus acetohydroxy acid synthase coding sequence.
 DE
 XX Large subunit; small subunit; Rhodobacter capsulatus;
 KM acetohydroxy acid synthase; AHAS; amino acid; carbon dioxide;
 KM photosynthetic microorganism; ds.
 XX Rhodobacter capsulatus.
 OS
 XX JP08214882-A.
 XX
 XX 27-AUG-1996.
 PD
 XX 16-FEB-1995; 95UP-00028522.
 PF
 XX 16-FEB-1995; 95UP-00028522.
 PR
 XX 16-FEB-1995; 95UP-00028522.
 XX
 XX (CHIK-) ZH CHIKU KANKO SANGYO GIJITSU KENKYU.
 PA
 XX WPI. 1996-436937/44.
 DR
 XX P-PSDB; AAW06554.
 XX
 XX Aceto:hydroxy acid synthase gene from Rhodobacter capsulatus - for
 PT improving amino acid production in photosynthetic microorganisms.
 XX
 XX Claim 2; Page 7-9; 10pp; Japanese.
 PS
 XX
 XX The sequences given in AAT45893-94 encode the large and small subunits of
 CC the Rhodobacter capsulatus acetohydroxy acid synthase (AHAS). This
 CC sequence is given in the claims as encoding the small subunit, and in the
 CC disclosure of the specification as encoding the large subunit. AHAS
 CC produces an amino acid using carbon dioxide as a carbon and energy
 CC source. Amino acid productivity can be increased by improvements in a
 CC photosynthetic microorganism through transformation with the AHAS gene
 XX
 SQ Sequence 1734 BP; 334 A; 525 C; 573 G; 302 T; 0 U; 0 Other;

Query Match 18.8%; Score 368.6; DB 2; Length 1734;
 Best Local Similarity 53.7%; Pred. No. 6.3e-106;
 Matches 937; Conservative 0; Mismatches 775; Indels 34; Gaps 7;

QY 136 ACCGGGCTTATATCTGATGAGTACCTGAAACGCGCATGGGTCACAAACATTTTGGC 195
 DB 4 ACCGGCGGGAATGCTGTGCTGACGCGCTGAAGATCAGGGGCTGACACATCTTCGA 63
 QY 196 TATCCCGCGGGGCAATTTTGGCCATCTATGATGAAGTACCGCTTTGAAGCGCGGG 255
 DB 64 TACCGGGGCGGCGCTGCTTCGATCTATGACGCGATCTTTCAGCAAAAC----- 114
 QY 256 GAAATTGACATATTTTGGTGGCCATGAACAAGAGCTTCCATGGCGGAGAT 315
 DB 115 GACATCCGCCACATCTCTGTGCGCATGAAGCGCGGTCACATGGCCGAGCTTAT 174
 QY 316 GCCAGAGCCACAGGTAAAGTGGAGTTTGTTCGTACATCTGACCAAGGCGCATAC 375
 DB 175 GCGCGCTGACCGGAAACCGGCGGTGTGTGCTGTGACCTCGGGCCCGGGGCGACAT 234
 QY 376 TTGGTGAACCGGATGCGCATTTGACATTTGACCTCGGCGCCATGCTGTGATTA 435
 DB 235 GCGGTGACCGGCGCTGACCGATGCTTGTGAGACAGCATTCGCTGTGCTTCTCGGT 294

QY 436 GAGGTGGCGCGTGCATGATTTGTAAGATGCTTTCCAGGAAATGACATTTTGGCATC 495
 DB 295 CAGGTGGCCACCTTCATGATCGGACCGACGAGTTTTCAGGAAGCCGACATCGGCATC 354
 QY 496 ACCTTACGATTCGTTAAGCACTCCTATGATGATGATGATGATGATGATGATGATGAT 555
 DB 355 ACCGCGCCCTGACACCAAGCACTGCTGATGATGATGATGATGATGATGATGATGAT 414
 QY 556 GTTACTGAGGCTTTCATCTGCTGACACCGGTCGTCGCGGCGGTTTGTATGATAT 615
 DB 415 ATCATCAGGCTTTCATGCTGACACCTCGGGGCGGCGGCGGCTTGTATGATGAT 474
 QY 616 CCCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 675
 DB 475 CCCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 534
 QY 676 CTACCGGATATGCGCCCAAGTTAAAGATATCCCGACCAATTAATGCGCATTCGCA 735
 DB 535 CAC-----TATACGCCCAAGTCAAGGGGATGATGATGATGATGATGATGATGAT 588
 QY 736 TTGTTGAGCAGGCGAAGAAATCCCTTGTCTTACGTAAGGGGAGGCGATGCGCCAT 795
 DB 589 CTGATCGAAGAGGCGGAAACCGCGATCTTTTACACGCGCGGCGGATCACTCAGCG 648
 QY 796 GCCCATGCCAGG-----TGAGGAAATTTGGGAAAGTTCCAGTTGCGGTAACAC 849
 DB 649 CCGCGGCGCTGCGCACTCTTGGCGAGCTGTGATGACAGCGGCTTCCGATCCTCG 708
 QY 850 ACCCTGATGGAATTTGGGCTTTTGAAGAAACATCCCTTTCGATGATGATGATGAT 909
 DB 709 ACCTGATGAGGCTGCGCGCTATCCGCTCGGCAAGGCGTGGCTCGCATGCTCGG 768
 QY 910 ATGCATGCGCCACCGCTATGCGCACTTTCGCTGAGGATGATGATGATGATGATGAT 969
 DB 769 ATGACGCGCTTTTACAGAGGCGCAATCTGGGATGACAGGCTGCTGATGATGATGAT 828
 QY 970 GGGGCGGCTTTCAGCAGCGGCTGACGCAATGAGCAATTTGCTGACCGCGCAAA 1029
 DB 829 GGGGCGGCTTTCAGCAGCGGCTGACGCAATGAGCAATTTGCTGACCGCGCAAA 888
 QY 1030 GTATTTCACTTATACATGACCGCGGAGTGGGAAAGAGGCTCCGATGTCG 1089
 DB 889 AAGCGCATGCTCATATGACGCTGCTGCTCATACCAAGATCATCTGCTGATCTGCGG 948
 QY 1090 ATTGTGGGGAATATACGCAATGTTTGAAGCTTTGACGCGGCGCGGAAATGAT 1149
 DB 949 ATGCTGGGATATCGGCGATGCTGAGGAGCTTCTCAAGGTCTGGAAGCGCGGAG 1008
 QY 1150 TACCCACCCATCCCATACACACCGAGGATGTTAAATGCAATGATGATGATGATGAT 1209
 DB 1009 CGCAAGTCAACAGGAGGCGTGGAGACTGATGAGGCGGATGACAGTGAAGAG 1068
 QY 1210 GATTACCCCTCCAGGTGCGCCACTATGAGATCAATTTGCCCCCAGAGATGATGAT 1269
 DB 1069 GTGCGCTGCTTCTCAACCACTGCGAGGATATCAACCGCAATATGCGCTGAG 1128
 QY 1270 GAAATTTGTGCGCAGGCG-----CCGATGCTTACTACACCAAGATGAGCAACAC 1323
 DB 1129 CGGCTGAGCGCTGACAGGAGGCGCAACCGCTATCATACCAAGMAATCGGCAAGAT 1188
 QY 1324 CAAATGAGGCGGCGCAATTTTGAACA--ATGCGCCCCCGCATGATGATTTCACTGCT 1380
 DB 1189 CAAATGAGGCGGCGCAATTTTGAACA--ATGCGCCCCCGCATGATGATTTCACTGCT 1248
 QY 1381 GCGTTGGTACATGAGGCTTTGATTTTACCTGCGCGCATGAGGAGCAAAATGAGAG 1440
 DB 1249 GCGTTGGTACATGAGGCTTTGATTTTACCTGCGCGCATGAGGAGCAAAATGAGAG 1308
 QY 1441 GAGGAGC-GCTCATTTGATCACTGAGATGACAGCTTCAATTAATTTTCAAGAACTG 1499
 DB 1309 GAGGAGC-GCTCATTTGATCACTGAGATGACAGCTTCAATTAATTTTCAAGAACTG 1368
 QY 1500 GGAACCTTACCGAGTACATCAAGTTAAACTATATTCTCAATTAAGCTTTGGCAG 1559

Db 370 ATTCATTAAGTTGTCCTTCCGCGCAGGTAGCGACTCGTTGATAGGTTACGATGCTTT 429
 Qy 472 CAGGAATTTGACATTTTGGCATCACCTTACCGATTCGTTAAGCACTCTTAAGGTACGT 531
 Db 430 CAGAGTCGCAATGATGAGGATTTCCGACCGGTGTTAACAAGTTTCTGATTAG 489
 Qy 532 AGTGGCGGATATGCTGCGATTTGTTACTGAGGCTTTCCATCTTGTACACCGGTCT 591
 Db 490 CAACGGGAAGACATTCGCGCAGGTCTGAAAAGGCTTCTGCTGCGCAAGTGGTGGC 549
 Qy 592 CCGGCGCGGTTTGTATGATATTCACAGGATGCGGCTTAGAAGATGTGATATT 651
 Db 550 CCAGGACGATGCTGATTTACCGAAGATATCTTAACTCCGCGAACAATTATCC 609
 Qy 652 CCGCTGACCCCGGTGACGTTAATCTACCGGTTATGCGCCACGTTAAAGTTATCC 711
 Db 610 TATGCTGCGGAGTGGTCAGTATGCTTCTTACATCCCACTACTACCGGACATPAA 669
 Qy 712 CGACAAATTAATGCGGCAATTGCAATTTGTTGAGCAGGCGAAGATCCCTTGCTTACGTA 771
 Db 670 GGGCAATTAAGCGTGTCTGCAACGCTGTAGCGGCAAAAAACGGTTGCTTACGTA 729
 Qy 772 GGGGAGGGGCGATGCGCGCAATGCGCATGCCAGGTGACAGAAATTTGCGAAAGGTT 831
 Db 730 GCGGTTGGGCGCAATACGCGCGGCTGCCATCAGCATTTGAAGAAACGTTGAGCGCTG 789
 Qy 832 CAGTTCCGCTAACACCACTGATGAGAAATGGGCTTTTGAAGAAACATCCCTT 891
 Db 790 AATCGCCGTTGTTGCTCATTTGATGGGCTGGGCGGCTTCCGCAACGCACTGCTAG 849
 Qy 892 TCGGTTGGTATGTTGGGTATGATGATGCGCACCGCTATGCGCACTTTCGCTAGCGAATG 951
 Db 850 GCATGGGCAATGCTGGGAATGACGCTACTACGAGCGCAATGACATGATGATACGCG 909
 Qy 952 GATTGTTGATTTGACGTGGGCGCGGTTTCACGACGCGGTAAGCTGCGCAACTAGACGAA 1011
 Db 910 GATGTGATTTTCGCGGTGGGTTACGATTTGATGACCGAAGCAACATCTGCGAAG 969
 Qy 1012 TTTGCTAGCCCGCAAAAGTAACTGACATGACATGACCGCGGCGAGGTGCGAAAAAC 1071
 Db 970 TACTCCCAATATGCACTGTTCTGCAATGCAATGATGATGCTTCTTCAATTTCTAAAC 1029
 Qy 1072 AGGCTCCCGATATGTCCTCATTTGTTGGGATGATGCGCATGTTTGAACAAGCTTTGCG 1131
 Db 1030 GTGACTGGGATATCCGATTGTGGGATGCTCGCCAGTCTCGAACAAATGCTTGA 1089
 Qy 1132 CGGCGCCGGAATTGGATTACCCCATCCATCCCATACCAAGCATGGTTAATGCG 1191
 Db 1090 CTCTTGTGCAAGAAATCGCCCATCAACCACTGATGATCGGACTGGTGCAGCAA 1149
 Qy 1192 ATTGATCATTTGCGGACCGATTACCCCTCCAGGTGCCCACTATGAGATATATTGCC 1251
 Db 1150 ATTGAACATGCGCGCTGCTGCAAGTCTGAATATGACACTCAAGTGAAGAAATTAA 1209
 Qy 1252 CCCGAGAGGTAGTACAGAAATTTGTGCGCAGGCGC--CGATGCTTACTACACCA 1308
 Db 1210 CCGCAGGCGGTATGAGACTCTTTGGCGGTTGACGAAGGAGACGCTTACGTGACGTCC 1269
 Qy 1309 GATGTGGGACACACCAATATGTTGGCGGCGC--AGTTTGAACAATGCGCCCGCGCA 1365
 Db 1270 GATGTGGGACACCAATGTTTGTGCACTTTATTCATTTGACAAACCGGTGCG 1329
 Qy 1366 TGGATTTCCAGTGTGCTTGGGTGCAATGGGCTTTGTTTACTGCGCGCATGGAAGCC 1425
 Db 1330 TGGATCAATTCGGTGGCTCGGCAAGTGGTTTGGTTTACCTGCGGCACTGGGCGT 1389
 Qy 1426 AAGTGGGAGTGGGAGACGAG--CGTCAATTTGCAATCAGTGAAGTCCAGCTTCCAATG 1484
 Db 1390 AAAATGGCTTCCAGAAAGAAACCGTGGTTGCTGCACTGGGAGCGCATTTCAAGT 1449
 Qy 1485 AATCTTCAGGAATCTGGAAACCTTACCGCATGACATCCAGTTAAATATTATTCTC 1544

Db 1450 AACATCCAGAACTGTTCTACCGCGCTTGCATACGAGTTCCCGCTACTGTGTGATCTC 1509
 Qy 1545 AATAACGTTTGGCAGGGATGTTGCTGCTGACGTCGCAACAACTTCTACGAAGAACATTAT 1604
 Db 1510 AATAACGCTATCTGGGATGTTGAAGCAGTGGCAGCAATGATCTATTCGCGCGCTCAT 1569
 Qy 1605 TCTGCTTCTAATG 1619
 Db 1570 TCACAACTTATATG 1584

Search completed: July 25, 2004, 05:27:58
 Job time : 830.718 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 25, 2004, 05:10:12 ; Search time 152.901 Seconds
(without alignments)
7110.132 Million cell updates/sec

Title: US-09-893-033-6
Perfect score: 1959
Sequence: 1 gccatagagcccatgcgcg.....cataagcccaaatgaattc 1959

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/ptodata/2/ina/5A.COMB.seq: *
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq: *
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq: *
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq: *
5: /cgn2_6/ptodata/2/ina/PCITUS.COMB.seq: *
6: /cgn2_6/ptodata/2/ina/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	416.6	21.3	1740	4	US-09-489-039A-5163
2	370.6	18.9	1716	4	US-09-252-991A-7372
3	370.6	18.9	1743	4	US-09-252-991A-7128
4	368.2	18.8	96109	4	US-09-596-002-35
5	366.8	18.7	4411529	3	US-09-103-840A-1
6	365.4	18.7	4403765	3	US-09-103-840A-2
7	353	18.0	1632	4	US-09-540-236-646
8	328.4	16.8	1830121	4	US-09-557-884-1
9	328.4	16.8	1830121	4	US-09-643-990A-1
10	323.8	16.5	1689	4	US-09-489-039A-840
11	321.6	16.4	1731	4	US-09-328-352-3942
12	318.8	16.3	2111	4	US-10-096-571-11
13	318.8	16.3	2111	4	US-10-096-571-13
14	309.2	15.8	2841	4	US-08-452-075-1
15	309.2	15.8	2841	3	US-09-231-061-1
16	309.2	15.8	2841	3	US-09-011-762-5
17	304.4	15.5	1947	4	US-09-489-039A-3367
18	304	15.5	6211	4	US-08-961-527-8
19	293	15.0	1969	1	US-07-737-851-1
20	293	15.0	1969	1	US-07-737-851-2
21	293	15.0	1969	1	US-07-894-062-1
22	293	15.0	1969	1	US-07-894-062-2
23	293	15.0	1969	4	US-09-096-562-1
24	293	15.0	1969	4	US-09-096-562-2
25	290.6	14.8	1728	1	US-08-403-866-12
26	290.6	14.8	2231	1	US-08-403-866-14
27	290.6	14.8	12720	1	US-08-403-866-11

28	288.2	14.7	1969	1	US-07-894-062-3	Sequence 3, Appl1
29	288.2	14.7	1969	4	US-09-096-562-3	Sequence 3, Appl1
30	286.6	14.6	1969	1	US-07-737-851-3	Sequence 3, Appl1
31	276.2	14.1	1020	1	US-09-252-991A-7248	Sequence 7248, Ap
32	275	14.0	1664976	4	US-08-916-421B-1	Sequence 1, Appl1
33	271.8	13.9	640681	4	US-09-790-988-1	Sequence 1594, Ap
34	271.6	13.9	1710	4	US-09-543-681A-1594	Sequence 1, Appl1
35	269.4	13.8	2156	1	US-08-321-356-3	Sequence 3, Appl1
36	269.4	13.8	2156	1	US-08-321-356-3	Sequence 1047, Ap
37	256.6	13.1	1755	4	US-09-134-001C-1047	Sequence 199, App
38	254	13.0	1689	4	US-09-543-681A-199	Sequence 1, Appl1
39	240.8	12.3	2365	3	US-08-363-208-1	Sequence 102, App
40	240.8	12.3	2365	3	US-09-137-478-1	Sequence 102, App
41	230.8	11.8	15249	4	US-08-956-171E-102	Sequence 11016, A
42	155.4	7.9	1869	4	US-09-252-991A-11016	Sequence 10943, A
43	155.4	7.9	1884	4	US-09-252-991A-10943	Sequence 11240, A
44	155.4	7.9	1959	4	US-09-252-991A-11240	Sequence 881, App
45	130	6.6	315	4	US-09-489-039A-881	

ALIGNMENTS

RESULT 1
US-09-489-039A-5163
Sequence 5163, Application US/09489039A
Patent No. 6610836
GENERAL INFORMATION:
APPLICANT: Gary Breton et. al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
TITLE OR INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 2709.2004001
CURRENT APPLICATION NUMBER: US/09/489, 039A
CURRENT FILING DATE: 2000-01-27
PRIOR APPLICATION NUMBER: US 60/117,747
PRIOR FILING DATE: 1999-01-29
NUMBER OF SEQ ID NOS: 14342
SEQ ID NO 5163
LENGTH: 1740
TYPE: DNA
ORGANISM: Klebsiella pneumoniae
US-09-489-039A-5163

Query Match	Score	DB 4;	Length	1740;
Best Local Similarity	55.8%	Pred. No. 8.8e-128;		
Matches	905;	Conservative	0;	Mismatches 699;
			Indels	19;
			Gaps	5;
QY	172	CATGGGGTCAACACATTTTGGCTATCCGGGGGCAATTTGGCCATCTATGATGA	231	
DB	64	CAGGGCTCAAGCAAGATTCGGCTATCCGGGGGCGAGTCTCATATCTATGATCG	123	
QY	232	CTGTACCCGCTTGAAGCGCGGGGAAATTGACATATTGTTGGTCCCATGAACA	291	
DB	124	TTACATACCCCTTG-----GCGGATCGATCATGTGTTGTCACGACAGCG	174	
QY	292	GCTTCCCATCGCGCGATGGGTATGCAAGCAGAGTAAATGGAGATTGTTCCGT	351	
DB	175	GCGGTGATATGGCGATGCTCGGGCGGCGACCGGGAGTCCGCTGTGCTGTG	234	
QY	352	ACATCTGAGACCAAGGCGACTTATGTTGACCGGATTTGCCAATGCCATTTGA	411	
DB	235	ACCTTGGGCTGGGGCGAGCAGACGCATCATGCGATTCGACACCTATATGAC	294	
QY	412	GTGCGCATGTGATGATTTATGAGAGTGGCGCGTCCATGATTTGATGCGAT	471	
DB	295	ATTTCGTTGATATCCCTCTGGGGCAGGTAGCCACTCGCTTATCGCTACATG	354	
QY	472	CAGAAATTTGACATTTTGGCATCACTTACCGATGTTAAGACATCTATGTAG	531	
DB	355	CAGAGTGCACATGTGGGATTTCCCGTCCGTGTGAAGACAGTTTCTGTTGA	414	
QY	532	AGTGGCGGATATGCTGCTGATTTTCTGAGGCTTTTCATTTGCTAGCACCG	591	

Db 415 CAGACCGAATATTCCTCCGCGCTGCTGAAGAAAGCTTTGCTGCGCGCCACGCGTCCG 474
 QY 592 CCCGCGCGCTTTTGTATGATATATCCAGAGATGCGCTTAGAAGATGATGATCAT 651
 Db 475 CCCGCTCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 534
 QY 652 CCCCTCGACCCCGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 711
 Db 535 TACGCTGCGCGAGCGAGTGAAGTGCCTGCTGATTAACCAACACCAAGTGGCATAA 594
 QY 712 CGACAAATTAATGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 771
 Db 595 GCGGAGATTAACCGCGCTGACAGCTGCTGCTGCGCGAGTAAGCGCGCTGCTGATG 654
 QY 772 GCGGAGAGGCGATGCGCGCAATGCGCGCAATGCGCGCAATGCGCGCAATGCGCGCAAT 831
 Db 655 GCGCGCGCGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 714
 QY 832 CAGTTGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 891
 Db 715 AAGCTGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 774
 QY 892 TCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 951
 Db 775 GCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 834
 QY 952 GATTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1011
 Db 835 GACGTGATTTTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 894
 QY 1012 TTTGCTAGCGCGCGCAAGTATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 1071
 Db 895 TATTGCTCGAAGCGCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 954
 QY 1072 AGGGCTCCCGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1131
 Db 955 GTGCGCGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1014
 QY 1132 CGGCGCGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1191
 Db 1015 CTGCTGAG 1074
 QY 1192 ATTGATCTTTGCGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1251
 Db 1075 ATCGAG 1134
 QY 1252 CCCGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1308
 Db 1135 CCGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1194
 QY 1309 GATGAG 1365
 Db 1195 GATGAG 1254
 QY 1366 TGGATTTCCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1425
 Db 1255 TGGATTTAACTCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1314
 QY 1426 AAAAGTGGAGTGGGAG 1484
 Db 1315 AAAAGTGGAGTGGGAG 1374
 QY 1485 AATCTTCAG 1544
 Db 1375 AACATTCAG 1434
 QY 1545 AATAAGTGGAG 1604
 Db 1435 AATAAGTGGAG 1494
 QY 1605 TCTGCTTTAAGATGCTGAG 1664
 Db 1495 TCGAGTCTATATG--GAATCGCTACCGAGCTTTGCTGCTGCGGAGAGAGAGAGAGAGAG 1551

QY 1665 ATCAAGGATTAATCTGCTGCGCAAGCGGAGAGATTTGCGCGCGGATGCGCGAATGCTA 1724
 Db 1552 CACGTGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGG 1611
 QY 1725 GCCCAATAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1784
 Db 1612 GAGCAGTGTGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGAT 1671
 QY 1785 ATG 1787
 Db 1672 GTG 1674
 RESULT 2
 US-09-252-991A-7372/c
 ; Sequence 7372; Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubinfeld et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; TITLE OF INVENTION: AERUINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; PRIOR FILING DATE: 1999-02-18
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 7372
 ; LENGTH: 1716
 ; TYPE: DNA
 ; ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-7372
 Query Match 18.9%; Score 370.6; DB 4; Length 1716;
 Best Local Similarity 54.5%; Pred. No. 1.8e-112;
 Matches 884; Conservative 0; Mismatches 714; Indels 25; Gaps 6;
 QY 135 AACCGGCGCTTATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 194
 Db 1714 ACTGCGCGTGAATAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1655
 QY 195 CTATCCCGCGGCGGCAATTTTGGCATCTGATGATGATGATGATGATGATGATGATGATGAT 254
 Db 1654 GTACCGGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1604
 QY 255 GGAATGAGCATATTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 314
 Db 1603 CAGCTGACCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1544
 QY 315 TCCCAAGCCAGAGTAAAGTGGAGTTTGTTCGTAATCTGATGATGATGATGATGATGATGAT 374
 Db 1543 CCGCGCGCGCACCGCGCAAGCCCGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1484
 QY 375 CTGGTGAACCGGATGCTGCAATGCGCATTTGGAATGCGGATGCGGATGCGGATGCGGATGCTG 434
 Db 1483 CCGCATACCGGATGCTGCAATGCGCATTTGGAATGCGGATGCGGATGCGGATGCGGATGCTG 1424
 QY 435 AGAGGTGGCGCTGCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 494
 Db 1423 CAGAGGTGGCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1364
 QY 495 CACCTTACCGATGCTGTAAGCACTCTATGCTGATGATGATGATGATGATGATGATGATGATGAT 554
 Db 1363 CTCCCGCGCGATGCTGTAAGCACTCTATGCTGATGATGATGATGATGATGATGATGATGATGAT 1304
 QY 555 TCTTACTAGGCTTTCACTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 614
 Db 1303 GATCAAGAGAGCTTTTACCTGCTGCGAGTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1244
 QY 615 TCCCAAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 674

Db	1243	TCCGAAGCATGGGCGACCCGACGACAGAAAGTTGCAATATTCCTATCCGAAGAAAGCTCA	1189
QY	675	TCTACCGGGTATACGCCCAACGGTTAAAGGTATATCCCGACAAATTAATGCGGCAATTGCA	734
Db	1183	GTTGCGTTGCAACAGCCCGCGCTTGCGGGTCACTGGAAGACAGATCCGCAAGCGCCGCA	1122
QY	735	ATTGTTGAGCAGGCGACAAAAATCCCTTGCTCTACGTAGGGGAGGGGAGATGCGCCGAA	794
Db	1123	GATGCTCTGGCCCGCAGAGCGCCGGTGTCTATTCGGCGGGCGCGTGATCATATGGGCA	1066
QY	795	TGCCCATGCCAGGTGACAGAAATTTGGGAAAAGTTCCAGTTGCCGTGTAACAACCTT	854
Db	1063	TGCCGCGCGCGCGCTGACCGAGCTGGCGACAGTGTCTCAACTGCGCGGTGACCMACCTT	1004
QY	855	GATGAGAAATGGGCGTTTGAAGAAAACATCCCTTGCGGGGTATGTTGGGTATGCA	914
Db	1003	GATGGGCTCGGGCGGGTATCCCGGGCGGACCGCAATTCCTTGSCATGCTCGGATGCA	944
QY	915	TGGCCACCGCTATGCGAACTTTGCCGTACGGGATGTATTTGTTGATTCAGTGGGAGC	974
Db	943	CGGCAAGCTTCAACCGCAACTGGCGATCATCAACGCACTGATCTCGCGGTGGGCGC	884
QY	975	CCGTTTGACACACCGGGTAACTGGGAAACTAGACAAATTTGGCTACCGCGCAAGTAAT	1033
Db	883	GCGTTTGACACACCGGGTATCAACGCGCGCGCAAGTTCTGCCGAACCGCAATCAT	824
QY	1035	TCACATTCATCAACGACCGCGGAGGTGGAAAAACAGGCTCCCGATGTGCCCATGT	1094
Db	823	CCACATGACATCATCGCCGGGTGATTTCAAGACATCAAGGCGCATGCCCATCGT	764
QY	1095	GGGGAGTGTACGCCATGTTTAAACACCTTTTGACGGGGCGGGGAATTGGATTACC	1155
Db	763	CGGCGCGGTGACAGCGTTTCTACCGGAATGTCGCGATCTGTAGGGAAATGGCGGAC	704
QY	1155	CACCCATCCCATACCAACCAAGCATGGTTAAATGCATTGATCATTTGAC---	1211
Db	703	CCCGAACAGAGATGCCAGGCTGCTGTGTGAAGACAGATCAACGATGGGAGTTAACG	644
QY	1212	TTACCCCCCAGGTGGCCCCCACTAAGAGATATCTATTCGCCCGCAGAGATGATACAGA	1271
Db	643	TGGGTTGTTCCCGTACGACAGAGGCGACGGCGATCATCAAGCGCGACCGGTATGGA	584
QY	1272	AATTGTCGACAGGCCCC-----CGATGCTTACACACACCGATGTGGGACACACA	1322
Db	583	GACGCTACAGAGGTATCCCATGGGAGTGCCTTATCATCTCGATGTGGGCAACCA	524
QY	1326	AATGTGGCGGGCCAGTTTGGAA--CAATGGCCCCCGCGATGGATTTCAATGCTGG	1387
Db	523	GATGTTGCGCTCGCAGTACTCAAGTTAAACAAGCCCAATCGTGGATCACTCGGGCG	464
QY	1383	CTTGGGTACAGTGGGCTTTGGTTTACTGCGCCCATGGAGCCAAAGTGGAGTGGGGA	1442
Db	463	TCTGGGACCAATGGGCTTCGGGTTCCGGCGGCGAATGGGCAATCAAGTCACTTCCGGA	404
QY	1443	CGA-GCGGTCAATTTCATCACTGAGAGATGCCAGCTTCCAAATGAATCTTCAGAACTGG	1501
Db	403	CGAAGATATGCTCGGTGACCGCGGAGGAGGAGATTCAGATGAATATCCAGAACTGTC	344
QY	1502	AACCTTAGCCAGTACGACATTCAGAGTTAAACTATTTCTCATTAACGTTGGCAGG	1561
Db	343	TACCTGCTGAGTACGATCGCTGGGTAAATGCTCACTGAACTAATGTCCTTGG	284
QY	1562	GATGTTGGGTCAATGGCAACAATTTCTAGAGAAAGTTATTTGCTTCTACATGTC	1621
Db	283	CATGCTCGCAGTGGCAGGACATGCAATGACAAACAGCGTTATTCCTATCATG--	226
QY	1622	CCAGGGCATGCGACAGATTAATCTCTCTGTGAAGCCTATGGCATCAAGGATATCTGT	1681
Db	225	-GAATCTGCGGACTTGCTGCAAGCTTGGCGAATCTTACGGCATGTGCGATCCGAT	167
QY	1682	GCGCAAGCGGAAAGTTTGGCCCGCGCATGCGGAAATGCTAGCCCAATGTCCTGT	1741

Db	166	CACCGACCTGAAGATCTGAAGCCGGAAGATGGAGAGGGCTTGCCATGAAGAATGCGCT	107
Qy	1742	GGT 1744	
Db	106	GGT 104	

RESULT 3
US-09-252-991A-7128
; Sequence 7128, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 7128
; LENGTH: 1743
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-7128

Query Match 18.9%; Score 370.6; DB 4; Length 1743;
Best Local Similarity 54.5%; Pred. No. 1.8e-112;
Matches 884; Conservative 0; Mismatches 714; Indels 25; Gaps 6;

Qy	135	AACCGGCGCTTATATCTCTGATGATAGCTGAACCGCATGGGGTCAACACATTTTGG	194
Db	30	ATCTGGCGCTGAAGATGCTGTCGCTCGTTGCCGACGAGGCGTTAAGTACATCTACGG	89
Qy	195	CTATCCCGGGGGGCGCAATTTTGGCCATCTATGATGAACGTGACCGCTTGAAGGGCGGG	254
Db	90	GTAACCGGGGTGGTCCCTCTGTCATATCTACGACGGGCTTTCTT-----AAGAGCA	140
Qy	255	GGAATTTGAGCATATTTTGGTGGCCATGAACAGAGAGCTTCCATGCGCGGATGGGTA	314
Db	141	CGACGTGACCCCAATCTCGGTGGCCACGACGAGCGGCTACCCCAATATGGCCACGGCTA	200
Qy	315	TGCAGAGCCACAGGTAAGTGGAGTTTGTTCGGTACATCTGAGACGAGGGCGACTAA	374
Db	201	CGCCCGGCCACCGGCAAGCCCGGTGTCTGTCTGTGATCTTCCGGCCCGGGCGGACCA	260
Qy	375	CTTGTGACCGGCAATTTGCCAATGCCAATTTGAATCGATGGTGGTGTGATTA	434
Db	261	CGGCATACACGGTATGCGCACCGCATATACATGACATCCATCCGATGGTATCTGTGGG	320
Qy	435	AGAGGTGGGCGCGTGCATGATTTGGTAGCATGCTTTCAGAGAAATGACATTTTGGCAT	494
Db	321	CGAGGTGCCGAGCAACATGCTCGGTACCGAGCGGTTTCAGAGAAACGACATGGTGGGAT	380
Qy	495	CACCTTACCGATGTTAAGCACTCTATGTGTAGTATGTCGGCGGATATGCTCGCAT	554
Db	381	CTCCCGGCGATGTAAGACACAGTTTCATCATCAAGCATCCCTCGGAATTTCCGAGGT	440
Qy	555	TGTTACTGAGGGCTTTCATCTTGTGTAGCACCGCGGTCTCCCGGGCGGTTTGTATCAT	614
Db	441	GATCAAGAGGCGCTTTTACCTGGCGAGTCCGCGTCTCGGGGCGGTTGTCTGTGATAT	500
Qy	615	TCCCAAGATGTGGGCTTGAAGAAATGTGATCATTTCCCTTGAACCCCGGTACGTTAA	674
Db	501	TCCGAAGACATGGGCGACCCCGACGAGAAAGTTGCAATATTTCTATCCGAAGAAGTCAA	560
Qy	675	TCTACCGGGTATTCGCCCGCACCGTTAAAGTATCCCGCAATTAATCCGGCATTTGCA	734
Db	561	GTTCGCTGTACAGCCCGCGGTTCGGGTACTCGGAGCAATCCGCAAGGCGCCGCA	620

QY 735 ATTGTGAGAGGCGAGAAATCCCTTGTCTACGTAGGGGAGGGCGATCGCCGCA 794
 Db 621 GATCTCTGCGCCCGAGCGCCCGGTGTCTATTCGCGGCGCGCGATGATCAAGGCA 680
 QY 795 TGCCCATGCCAGGTGCGAGATTTGCGAAAAGTTTCAAGTTGCGGTTAAACAACCT 854
 Db 681 TGCCGCGCGCGCGGTGACCGAGCTGCGGCAAGTCTCACTCGCGGTGACCAACCT 740
 QY 855 GATGAGAAATTTGGCTTTTGAAGAAACATCCCTTTGCGTGGATGTTGGATGCA 914
 Db 741 GATGAGCTCTCGCGGAGATCCCGCGGAGACCGCGCACTTCTTGGCATGCTCGGATGCA 800
 QY 915 TGCGACACCGGTATGCCAATTTGCGTACAGCAATGATTTGATGATGAGAGGCG 974
 Db 801 CGGACGCTTCAACCGCAACCTGCGGATGATCAAGGCACTGATCTCTCGCGGTGCGCG 860
 QY 975 CGGTTGACGAGCGGGTACTGAGCAAACTAGACAAATTTGTAACCGCGCCAAAGTAAT 1034
 Db 861 GCGTTGACGAGCGGGTACTGAGCAAACTAGACAAATTTGTAACCGCGCCAAAGTAAT 920
 QY 1035 TCACATGACATGACACCGCGGAGGTGAGAAAACAGAGGCTCCGATGTGCCATGTT 1094
 Db 921 CCACATGACATGACATCGCGGCTGATTTCCAAACATCAAGGCGACATCCGATGCT 980
 QY 1095 GGGGAGTGTACCGCATGTTTGAACAGCTTTGACAGCGGCGCGGGAATGATTAACC 1154
 Db 981 CGGCGCGGTGACAGCGTTCTACCGAAATGTCGATGTCAGGAAATCGCGCGAGAC 1040
 QY 1155 CACCATTCCTTACACCGCAGGATGTTAAATGATGATGATGTC---GACCGA 1211
 Db 1041 CCGGACCGAGATCCGAGCGCTGCTGAGAACAGATGACAGAGGCGTGTACCG 1100
 QY 1212 TTACCCCTCAGGTGCGCCACTATGAGATATGTTGCCCCAGAGGTAGTACGA 1271
 Db 1101 TGGGTGTTCCGTAACGACAGAGGCGAGAGATGATCAAGCGGACGACGATGCA 1160
 QY 1272 AATTGTCGCGAGGCCCC---CGATCTTACTACACCGATGTGGACACCA 1325
 Db 1161 GACGCTTACAGGTTTCCAGTGGCGATGCTTATCACTCCGATTCGCGCAGACCA 1220
 QY 1326 AATGCGCGCGCCAGTTTGA---CAATGCCCCCGCGAGATGATTTCCAGTGTG 1382
 Db 1221 GATGTCCTCTGAGTACTCAAGTTCAACAGCCCAATGCTGATCACTCGCGG 1280
 QY 1383 CTGAGTACGATGCGCTTTGTTTACCTGCGCCAGTGGAGCCAAAGTGGAGTGGGA 1442
 Db 1281 TCTCGGACAGAGGCTTGGCTTCCGCGCGGATGGGATCAAGCTCAACTCCGGA 1340
 QY 1443 CGA-GCGGTCAATTGTCATGAGAGATGCGAGTCCAAATGATTTGAGGACGCG 1501
 Db 1341 CGACATGTGCTGCGTGGAGCGGAGGAGATGCAAGTGAATGACAGGACTGTC 1400
 QY 1502 AACCTAGCCAGTACGATCAGTAAATTAATTTCTCAATTAAGCTTGGCAGG 1561
 Db 1401 TACGTGCTGACATGACGCTGCGGTGAAGATGTCACCTGAACATGTGCGCGG 1460
 QY 1562 GATGTGCTGATGAGGCAAACTTTCTAGAAAGCTTATTTCTCTTAACATGTC 1621
 Db 1461 CATGTGCGCAGGAGGAGGAGATGACATCAAGCGCTTATTCGATTCACAG-- 1518
 QY 1622 CGAGGCGATGCGACATTAATCTCTGTGAAGCTTATGCGATCAAGGTAATTAATG 1681
 Db 1519 -GATTCCTGCGAGCTTCTGCAAGCTTCCGAGGCTTACGCGATGCGGATGCGCAT 1577
 QY 1682 GCGCAGCGGAGAAATTTGCGCGCGCGGATGCGGAAATGCTAGCCCAATGTCTGT 1741
 Db 1578 CACGACCTGAGAGTGAAGCGAAGATGAGAGGCGCTTCCGCAATGAAGATCGCT 1637
 QY 1742 GGT 1744
 Db 1638 GGT 1640

RESULT 4
 US-09-596-002-35
 ; Sequence 35, Application US/09596002
 ; Patent No. 6632636
 ; GENERAL INFORMATION:
 ; APPLICANT: Lagace, Robert, E.
 ; APPLICANT: Paterson, Chandra
 ; APPLICANT: Berg, Kim L.
 ; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF MORAXELLA CATARRHALIS GENOME
 ; FILE REFERENCE: PM-008-4 US
 ; CURRENT APPLICATION NUMBER: US/09/596,002
 ; PRIOR FILING DATE: 2000-06-16
 ; PRIOR APPLICATION NUMBER: 60/140,121
 ; NUMBER OF SEQ ID NOS: 41
 ; SOFTWARE: PERL Program
 ; SEQ ID NO 35
 ; LENGTH: 96109
 ; TYPE: DNA
 ; ORGANISM: M. catarrhalis
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; OTHER INFORMATION: Inocyte template ID No. 6632636 35
 ; PUBLIC INFORMATION:
 ; US-09-596-002-35

Query Match 18.8% Score 368.2; DB 4; Length 96109;
 Best Local Similarity 53.9% Pred. No. 1.7e-110;
 Matches 908; Conservative 0; Mismatches 743; Indels 34; Gaps 6;

QY 137 CCGGAGCTTATATCTGATGATGAGTGAAGCCCTGAAGCCATGGGTTAAACATTTTGGCT 196
 Db 2433 CAGGGGCTGAGATGCTTGTTCAGCATGATGATGAGAGGTGAGATATTTTGGCT 2492
 QY 197 ATCCGCGGGGCAATTTTGGCCATCTATGATGAACTGACCGCTTTGAACCGCGGGG 256
 Db 2493 ATCCGCGGGGTCGGATTTATCATATTTATGATGCTTTATTAACAGATAT----- 2543
 QY 257 AATTGACATATTTGTGTCGCGCATGACAGAGAGCTTCCATGCGCGGATGAGTATG 316
 Db 2544 AATTGACATATTTGTGTCGCGCATGACAGAGAGCTTCCATGCGCGGATGAGTATG 2603
 QY 317 CCAGAGCCAGAGTAAATGAGAGTGTTCGATGACATCGAACGAGGCGGACTACT 376
 Db 2604 GTGTGACAAAGCGAGCGGCGGCTGTTTGGCGACTTCAGGTCAAGTGCACAAACA 2663
 QY 377 TGGTACCGGATGTCGATGCGCATTTGATGCTGCGTCCCATGCTGATGATGATGAG 436
 Db 2664 CTGTAAAGCATTTGCACTGCTTATATGATGATTCATCCATGATGTTGTTGCGAGGTC 2723
 QY 437 AGGTGGCGCGTCCATGATTTGATGAGGATGCTTTCCAGAAATGACATTTTGGATCA 496
 Db 2724 AGGTGACATTAATTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2783
 QY 497 CCTTACGATGCTTGAAGACTCTATGATGATGATGATGATGATGATGATGATGATGATG 556
 Db 2784 CAGCGCCCATGCTTGAAGCAAGTTTCAAGTTGCGCAAGCGGAGAGATTCACAAATCA 2843
 QY 557 TTACTGAGGCTTTCATCTTGTGACACCGGTGCTCCGCGCGGCTTTGATGATGATTC 616
 Db 2844 TCAAAAAGCATTTTATATGAGGCTGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2903
 QY 617 CCAAGAGTGGGCTTAAAGATGATGATGATGATGATGATGATGATGATGATGATGATG 676
 Db 2904 CAAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2963
 QY 677 TACCGGTTATGCGCCCAAGTTAAAGTAAATCCCGCAAAATTAATGCGGATGCAAT 736
 Db 2964 TCAGATCATATGAGCTTACATTAAGTCAATGCGGAGCAATCAAAAAGCATCGAAA 3023
 QY 737 TGTGAGAGCGCCAGAAATCCCTGCTCTAGTACGATGAGGAGGAGGCGCATGCCCGCATG 796
 Db 3024 GGTACTATCAAGCCAAAGCGCTTATCTATGATGAGGAGGAGGAGGAGGAGGAGGAGG 3083

QY 797 CCCATGCCAGGTGACAGAAATTTGGGAAAAGTTCCAGTTGCGGTAACACCCCTGA 856
DB 3084 CCAGTAAGAGAGCTGGGTCAATTAGCCGATACGCTCAACCTCTTACTAATACCTTA 3143
QY 857 TGGGAAATTTGGGCTTTTGAAGAAAACATCCCTTGGGTGATGTTGGGTATGCATG 916
DB 3144 TGGGAACTGGGCTCTTATCCAGGCTCAGGCGAACAATTTGGAATGCTTGGTATGCATG 3203
QY 917 GCCACGCGTATGCCAACTTGGCGTACAGGAATGATTTGTTATGTCAGAGGGGGGCC 976
DB 3204 GCACATATGAGAGCAATGACATGATCATGCTGATGATGATTTGGCAGGGGTGTC 3263
QY 977 GTTTCAGAGACCGGTAACCTGAGCAAACTAGACGAATTTGCTAGCCGCCCAAGTATTC 1036
DB 3264 GTTTTGAATGACCGGTGACCAATTAATGCAAAAATTTGCCCCAATGCAACGATCATTC 3323
QY 1037 ACATTCAGATGACCCCGGAGGTGGGAAAAAAGGGCTCCGATGCTCCCATTTGTCG 1036
DB 3324 ATATTGATATGATCCAGCAGATCTCAAGACGATTAATGACACATTCGATGTTG 3383
QY 1097 GCGATGTAACGCGATTTTGAACAGCTTTTGCAGCGGCGCGGAAATGGATTAACCCA 1156
DB 3384 GCGATGTAAGCTGTGATCAATTAATGAAATGCTTGGCTTCTGACAGAGATTTGATGTA 3443
QY 1157 CCCATCCCATTCACCAACGCGATGTTAATCGCATGATCATTTGCGGAC----- 1208
DB 3444 TCGATGAACATGCGCTCAATGATTTGTGGCAACAAATCAATGAATGGGTAAACGCCATG 3503
QY 1209 -----GATTAACCCCTCCAGGTGCGCCACTAGAGGATTAATTTGCCCCAGAGG 1261
DB 3504 GTCTTCGCTATATGACATATATGATGCTCCAAACATGCTCAAGCCCAAGCGATG 3563
QY 1262 TAGTACAGAAATTTGTCGCCAGG---CCCGATGCTTACTACCAACCGATGTCGAC 1318
DB 3564 TGGTGAAGTGTCTTATTAAGCTGACGAATGGTGAACATATTAATCTCCATGTTGTC 3623
QY 1319 AACACCAATGTTGGGCGGCCAGTTT---TGAACATGGCCCCCGCGAGATTTCCA 1375
DB 3624 AGCATCAAGTGTCTGATTTGATTAATTAAGTATGATGACCAACGCAATGTTAACT 3663
QY 1376 GTGCTGGCTTGGGATGATGAGCTTTGCTTTTACCTGCGCATGAGGACCAAGTGGAG 1435
DB 3684 CAGGTGGCTTAGGTACATAGGCGCTTGGTTTACCTTAGCTATGACAGCAAAATTAAGTC 3743
QY 1436 TGGGGGAGAG--CGGTCAATTGTCATCATGAGATGCCAGCTTCCAAATGAATCTTCAAG 1494
DB 3744 ACCCAAAAAGACATGTTGTTGATACCTGGGAGAGCTCAATCAATGAATCAACAG 3803
QY 1495 AACTGGGAACCTTACCCAGTACGACATCCAGTTAAACTATTCTCAATAACGCTT 1554
DB 3804 AGCTGTCACTTGTCTACAGTAACTATCAACAGTTAAATTTAAATTAATTAATGCTC 3863
QY 1555 GGCAGGGAGTGTGCTGATGAGCAAACTTTCTACGAAGAACTTATTTCTCTTCTA 1614
DB 3864 AGCTTGGCATGTTAAGCAGTGGCAGATATGCTTATGAAGTGCAGCATCTCAATCTT 3923
QY 1615 ACATGTCCAGGGAGTGCAGACATTAATCTCTCTGGAAGCCATAGCATCAAGGTA 1674
DB 3924 ATATG---CAATCTTACTCTGATTTTGTAAATTTGGCAGAAAGTTAAGGATTAAGGTCG 3960
QY 1675 TTACTGTGCGAAGCGGGAAGATTTGGCCCGCGATGCCGAATGCTAGCCCAATG 1734
DB 3981 TAAATATTAACAATCTCTCAACATGCAAGAAAGCTTAAACAGCCCTTAAGATGATG 4040
QY 1735 GTTCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1794
DB 4041 GTTGGTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 4100
QY 1795 CCGGC 1799
DB 4101 CTGGC 4105

RESULT 5
US-09-103-840A-1/C
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patemlin Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
US-09-103-840A-1

Query Match 18.7%; Score 366.8; DB 3; Length 4411529;
Best Local Similarity 53.3%; Pred. No. 7e-109;
Matches 944; Conservative 0; Mismatches 772; Indels 34; Gaps 8;

QY 88 CCTAATCCGAACAAATTTCTGAAACTGTTCTTACGCAACGCAACCGGGCTTAT 147
DB 3362917 CCAAGACCCGGCGGCTGCTCGAACAATGTCACGTGAGAGCTTACCGGTGACAG 3362858
QY 148 ATCCTGATGATAGCTGTGAACCGCATGGGGTCAAAACATTTTGGCTATCCCGCGGG 207
DB 3362857 GCGGTATCCGGTCCGTGGAGGAACTCGGGGTGATGATTTCCGGATTCGGGGCGGT 3362798
QY 208 GCATTTTGGCATCTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 267
DB 3362797 GCGGTGTGCGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 3362747
QY 268 ATTTGATGGCCATTAACAGAGAGCTTCCATGCGCGGATGGATGATGATGATGATGAT 327
DB 3362746 GTGCTGTGCGGCACAGAGAGGCGCGCGGATGCGGATGCGGATGCGGATGCGGAT 3362687
QY 328 GTTAAGTGGAGTTGTTTGGTACATCTGACAGGGGCGACTTAATCTTGTGACCGGC 387
DB 3362686 GGCCTGGGTGGCGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 3362627
QY 388 ATTGCAATGGCCATTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 447
DB 3362626 CTGACCGAGCGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3362567
QY 448 GCCATATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 507
DB 3362566 GGGCTGATGGGACCGACGCTTCCAGAGGCGGATCATCTCGGGATCAACATGCGGATC 3362507
QY 508 GTTAAGCATCTCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 567
DB 3362506 ACCAAGCACAATTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3362447
QY 568 TTCCATCTTGTAGCAACCGGTGCTCCGGGCGGTTTGTATGATGATTTCCCAAGATGTCG 627
DB 3362446 TTCCATGATGCGGCGCTCCGGCGGCTCGGGCGGCTGCTGCTGCTGCTGCTGCTGCT 3362387
QY 628 GCGTTAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 687
DB 3362386 CT---GCAAGGCGCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3362333
QY 688 CGCCCAAGGTTAAAGTAAATCCCGGACAAATTAATGCGGATGATGATGATGATGATGATGAT 747
DB 3362332 AAGCCCAACCAACCAACCGGACAGGCTGCGGAGGCGGACCAAGTATGCGGCGC 3362273
QY 748 GCCAGAAATCCTTGTCTTACGTAAGGAGGAGGCGGATGCGGCAATGCCATGCCAG 807

Db 3362272 GCGGCGAAGCCGCTGCTGTATGTGCGCGCGGCTCATCCGCGGTGAGGCGACCGAGCAG 3362213
 Qy 808 GTCCAGGAATTTGCGGAAAAGTTTCAGTTGCGCGTAAACAACACCTGATGAGGAATTGGG 867
 Db 3362212 CTCGGGAGCGTGGCGAGCTGACCGGCGATCCCGGTGCTACCAACGCTGATGCGCCGGGCG 3362153
 Qy 868 GCTTTTACGAAAACCATCCCTCTTCCGTGGGTATGTTGGGTATGCTAGCCGCTAT 927
 Db 3362152 GCGTTTCCCGACACCGACCGGCAAACTCGGCAATGCGCGCATGACGCGACGCTGGCC 3362093
 Qy 928 GCCAATTTGCGCGTACAGGAATGATTTGTTGATTCAGTTGGGGGCGCTTTCAGACAG 987
 Db 3362092 GCCGTGGGGGCGCTGCGACCGGACGACCTGCTGATCCCGTGGTACCCGCTTTCAGACAG 3362033
 Qy 988 CGGTTACTGSCAAACTAGACGAATTTGCTAGCCGCGCCAAAGTAAATTCACATTGACATC 1047
 Db 3362032 CGGGTGAACCGGCAAGCTGCACTCGTTCGCGCGCGGAAACCAAGGTCATCCACGCGCAATC 3361973
 Qy 1048 GACCCGCGGAGGAGTGGGAAAAAACAAGGCTCCGATGTCGCATTTGCGGGGATGTAACG 1107
 Db 3361972 GACCCGCGCGGAGATCGGAGAAACCGGCAACCGGACGCTGCCATCTGCTGAGCTCAAG 3361913
 Qy 1108 CATGTTTGAACAGCTTTTTCAGCGGCGCGGGAATGGAATTAACCCACCCATCCCAT 1167
 Db 3361912 GCGGTCAATCAACGAATGATGCGATGCTGCGCCACACACACATTCGCGGACCATGAG 3361853
 Qy 1168 ACCACCCGAGGATGTTAAATGCGATGATGATTTGGGGGACCGCATACCCCTCCAGATG 1227
 Db 3361852 ATGCGCGAGCTGTGGGCACTTCTGAAAGCTGTGCGCAAGCTTATCCGCTGAGCTATGG 3361793
 Qy 1228 CCCCATGTAGAGTACTATTGCCCCCAGGAGGTAGTATACGAAATTTGCTGCGCAGGCC 1287
 Db 3361792 CCGCAGAGCGAGCGGAGCTGAGCCCGGAATGATGATGCAAAAGCTCGGCGAGATGGCC 3361733
 Qy 1288 ---CCCCATGCTTACTACCAACCGATGTGGGACAAACCAATGTGGGGGCGCCAGTTT 1344
 Db 3361732 GAGCGGAGCGCGCTTGTGTGCGCGGCTCGCGGACGACAGATGTGGGCGCGCACTTC 3361673
 Qy 1345 TTGA---ACAATGGCCCCCGCGCAGATTTCCAGTCTGCTGCTGCTGATGAGTGGCTTT 1401
 Db 3361672 ATCAGATACGAAGACCGCGAGCTGCTGTAACCTCGGCGGTCTGCGCACCATGAGGCTTT 3361613
 Qy 1402 GATTACTGCGCGCGATGAGAGCCAAAGTGGAGT---GGGGAGCAGCGCTCATTTTGCATC 1460
 Db 3361612 GCATATCCGCGCGCGCATGGGCGCAAGATCGCTCCCGGACCGGAGTCTGGGCGATC 3361553
 Qy 1461 AGTGAAGATGCGAGCTTCAATGATCTTCAAGAACTGGGAACCTTAGCCAGTACAGC 1520
 Db 3361552 GACGCGAGCGGTGCTTTCAGATGACCAACGAGAGCTGCGCACCTGCGCGGTGAGGCGC 3361493
 Qy 1521 ATCCAGTTAAACTATTATTTCATATACGTTGGGACGCGGATGTGCTCACTGAGCA 1580
 Db 3361492 ATACCGGTCAAGGTGGCGCTGATCAACACGCAACTGGGCAATGTGGCGAGTGGCAG 3361433
 Qy 1581 CAACCTTCTACGAAGAACCTTATTTGCTTCTTAACATG-----TCCAGGGGATG 1631
 Db 3361432 AGCCTGTTCTATGCGAGCGGTACTGCGACGACCTGCGCACTCATTTGACCGGATC 3361373
 Qy 1632 CCAGCATTTAATCTCTCTGTGAAGCCTATGCGATCAAGGTTATTACTGTGCGCAACGG 1691
 Db 3361372 CCGGATCTCGGAACCTGCGGAGGCGCTTGGGATGTGGGCTTGGCGTGGAGCGGAA 3361313
 Qy 1692 GAAATTTGGCCCCCGGCGATGCGGGAATGCTAGC---CCACATGCTCTGTGTGATG 1748
 Db 3361312 GAGGAGCTGCTGACAGTCAACACGCGCGGCGATCAACGATCCCGGTGTGATC 3361253
 Qy 1749 GATGTGTGTCAAAAAGATGAAAACCTTACCTTATGATGTCGCCCGGATAGTAAT 1808
 Db 3361252 GACTTACCTGTGCGGCGGACGCGCAAGTGTGCGGATGTGGCGCGGCGGACCAAGCAAT 3361193
 Qy 1809 GCCCAATGC 1818

Db 3361192 GACGAGATCC 3361183
 RESULT 6
 US-09-103-840A-2/c
 ; Sequence 2, Application US/09103840A
 ; Patent No. 6294328
 ; GENERAL INFORMATION:
 ; APPLICANT: FLEISCHMAN, Robert D.
 ; APPLICANT: WHITE, Owen R.
 ; APPLICANT: FRASER, Claire M.
 ; APPLICANT: VENTER, John C.
 ; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
 ; FILE REFERENCE: 24366-20007 00
 ; CURRENT APPLICATION NUMBER: US/09/103,840A
 ; NUMBER OF SEQ ID NOS: 2
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 2
 ; LENGTH: 4403765
 ; TYPE: DNA
 ; ORGANISM: Mycobacterium tuberculosis
 ; FEATURE:
 ; OTHER INFORMATION: CDC 1551
 ; OTHER INFORMATION: "n" bases at various positions throughout the sequence
 ; OTHER INFORMATION: represent a, t, c or g
 US-09-103-840A-2
 Query Match 18.7% Score 365.4; DB 3; Length 4403765;
 Best Local Similarity 53.8%; Pred. No. 2.1e-108;
 Matches 942; Conservative 1; Mismatches 773; Indels 34; Gaps 8;
 Qy 88 CTTATCCCAAAACAAATTTCTGAAACATGTTCTTACCAACGAGCAACCGGAGCTTAT 147
 Db 3357249 CCAAGACACCGCGCGCTGCTCGAAACATGTGACATGACGACGATTAACGAGTACAG 3357190
 Qy 148 ATCTGATGATGATGCTTGAAGCCATGGGTCAACACATTTTGGCTATCCCGCGCG 207
 Db 3357189 GCGGTATCCCGCTGCTGAGGAACTCGGCGTCAAGCTGATTTTGGATTTCCGGCGGT 3357130
 Qy 208 GCATTTTGGCCATCTATGATGAACTGATGACCTGATGACGCTTGAAGCGCGGGAATTTAGCAT 267
 Db 3357129 GCGGTGCTGCGGTGATGACCCGCTGTTTCACTGAAAGATGCTGCGC-----CAC 3357079
 Qy 268 ATTTTGTGCGCGATGAACAAAGAGCTTCCATGCGCGGATGGGTATGCCAGACCA 327
 Db 3357078 GTGCTGTCCGCGCAAGAACAGGCGCGGATGCGCGCTACGCGCACGTCAC 3357019
 Qy 328 GGTAAAGTGGAGATTTGTTGATGATCTGGAACGAGGAGGAGATTAATTTGATGACCGGC 387
 Db 3357018 GCGCGGATGGCGGTGATGACGACGATGCGGTGCGGAGACCAACCTGTGACCCCG 3356959
 Qy 388 ATTCGCAATGCCCATTTGAGTCTGAGTCCCATGATGATGATTAATCTGAAGAGTGGCGCT 447
 Db 3356958 CTGGCGAGCGCGCATGAGTCACTCCGCGGTGCTATCAACGATGCGGCGG 3356899
 Qy 448 GCATGATTTGGAGGATGCTTTCCAGGAATTTGATTTTGGATCACTTCCCATC 507
 Db 3356898 GGGCTGATGCGGACGAGCGCTTCCAGAGGCGGACATCTGGGATCAACATGCGCATC 3356839
 Qy 508 GTTAAGCACTCTATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 567
 Db 3356838 ACCAAGCAACATTTCTGTGCTGCGCTGCGGTGACGACATTTCCGCGGTCTGAGCGAGCC 3356779
 Qy 568 TTCCATCTTGTGAGACGAGTGTCTCCGCGCGGCTTTTGAATGATATTTCCAAAGATGTG 627
 Db 3356778 TTCCATCTGCGCGCTTCCGCGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCG 3356719
 Qy 628 GGTATGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 687
 Db 3356718 CT-----GAGAGCGCGATGACGTTTCACTGCGCGCGGATG--GAGCTGCGCGCTAC 3356665

QY	688	CGCCCCACGGTTAAAGTATATCCCGACAAATTAATGCGGCAATTGATGTTGGACAG	747
Db	3356664	AAGCCCAACACCAAAACCGCACAGCCGGAGGATCCGCGAGGCGCGCAAGCTGATGCGCCG	3356605
QY	748	GCCAGAAATCCCTTGCTCTACGTAGGAGGAGAGGCGATCGCGCAATGCCATGCCAG	807
Db	3356604	GCGGCAAGCCGGATGTCTATGTGCGCGGCGGTCAATCCGCGGTAGAGCCACGACAG	3356545
QY	808	GTGCAGAAATTTGCGGAAAGSTTCCAGTTGCCGGTAAACACACCCTTGATGGAAATTGGG	867
Db	3356544	CTCCGGAGCTGCGCGAGCTGACCCGCGATCCCGGTGTCAACGACGTGATGACCCCGCGG	3356485
QY	868	GCTTTTACGAAAACCATCCCTTTCCGTTGGGTATGTTGGGTATGCAATGGCAGCCGTAT	927
Db	3356424	GCCGTGCGGCGCTGCAGCGAGGACCTTGATATCGCGTGGTACGATCCGCTTGACGAC	3356365
QY	928	GCCAACTTTGCCGTACGCGAATGTGATTGTGATTGCAAGTGGGAGGCCGTTTGCAGAC	987
Db	3356424	GCCGTGCGGCGCTGCAGCGAGGACCTTGATATCGCGTGGTACGATCCGCTTGACGAC	3356365
QY	988	CGGGTAATGGCAAACTTGAAGAAATTGCTAGCCGCGCAAGTAATTCACATTGACATC	1047
Db	3356364	CGGGTACCGGGAAGCTCGACTCGTGCAGCGAGAGCCAAAGTATCCAGCCGACATC	3356305
QY	1048	GACCCGCGAGTGTGGAAAAACAGGAGCTCCCATGTGCCATTTGTGGGGAATGTACG	1107
Db	3356304	GACCCGCGCGAGATCGCAAGAACCGGACGCGCAGCGATCGTCCGTGAGAGTGAAG	3356245
QY	1108	CATGTTTAAAGACGCTTTTGCAGCGGCGCCGGAAATTGAGTTACCCACCCATCCCAT	1167
Db	3356244	GCCGTATCACCGCAACTGATGCGATGTGTGCCACACACCATTTCCCGGACCATGAG	3356185
QY	1168	ACCACCCAGGATGGTAAATCGATGATCATTTGGCGGACCGATTACCCCTCCAGAGT	1227
Db	3356184	ATGGCCGACGTGTGGGCACTACTTGAAGGTGTGGCGAAGACTTATCCGCTGAGCTATGG	3356125
QY	1228	CCCCACTATGAGATACTATTGGCCCCCGAGAGGTAGTACAGAAATTTGTTGCGCAGGC	1287
Db	3356124	CCGCGAGCGCAGCGCAGCTGAGCCCGGAATACCTGATCGAAAAGCTCGCGAGATCGCC	3356065
QY	1288	---CCGATGCTACTTACACCAACCGAGTGGGAGAAACAAATGTTGGGCGCCAGTT	1344
Db	3356064	GGGCCGAGCGCTCTTCTGTGTGGCGGTGCGGCACGACCAAGATGTGGGCGCGAGTTG	3356005
QY	1345	TTGA--ACAATGGCCCCCGCGCATGTGATTTCCAGTCTGGCTTGGTACATGGGCTTT	1401
Db	3356004	ATCAGATTCGAAAAGCGCGCAGCTGGGTGAATCCGGCGGTCTGGGCACTAGGGTTT	3355945
QY	1402	GTTTACTCTGCCGCAATGGAGCCAAAGTGGAGAT--GGGGAGAGCGGCTCATTTGCATC	1460
Db	3355944	GCAATCCCGGCGCATAGGCGGCCAAATGCGCTCCCGCGCACGAGGTCTGGCGCATC	3355885
QY	1461	AGTGAAGATGCACACTTCCAATGAATCTTCAGGAATCGGAAACCTTAGGCCAGTACGAC	1520
Db	3355884	GAGGCGAGCGGTTGCTTCCAGATGACCAACGAGAGGTGCGCACTGCGGCTCGAGAGGC	3355825
QY	1521	ATCCAGGTTAAACTATTAATTTCTCAATAAAGGTTGGCAGGGAGTGTGCTCATGTGCCAA	1580
Db	3355824	ATACCGGTCAAGGTGGCGCTGATATCAACAGGCAACTTGGGCAATGTGGCGCGAGTGCAG	3355765
QY	1581	CAAACTTTCTACGAAGAAGCTTATTTCTTCAATG-----TCCGAGGGGATG	1631
Db	3355764	AGCTGTCTTATATGCCAGCGGTATTCGACAGACGACCTGGCGCACTATTTGCACCGCATC	3355705
QY	1632	CCAGACATTAATCTCTCTGTGAAGCCTATGGCATCAAGGTTATTACTGTGCGCAACCG	1691
Db	3355704	CCCACACTTCGAAACATGCGCGAGGCTTTGGGGTGTGTGGGGTTGCGGTCCGAGCGGAA	3355645
QY	1692	GAAGATTTGGCCCCCGGCGATGCGCGGAATGTGAC---CCACATGGTCTCTGTGGATG	1748
Db	3355644	GAGAGCTCGTCAACGTCATTAACGAGCGCGGCGCATCAACGCTGCCCGGTGGTGTATC	3355585
QY	1749	GATGTGGTGTCAAAAAAGATGAACCTGTTACCTTATGATTGGCCCCCGCATGAGTAAT	1808

Db	3355584	GACTTCACGTCGGTCCGACGCGCAAGTGTGGCCGATGTGGCCGCGGCGACACGACAT	3355525
Qy	1809	GCCCAATGC	1818
Db	3355524	GACGAGATCC	3355515

RESULT 7
US-09-540-236-646
; Sequence 646, Application US/09540236
; Patent No. 6673910
GENERAL INFORMATION:
APPLICANT: Gary L. Bireton et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CAT
FILE REFERENCE: 2709, 2005-001
CURRENT FILING DATE: 2000-04-04
NUMBER OF SEQ ID NOS: 3840
SEQ ID NO 646
LENGTH: 1632
TYPE: DNA
ORGANISM: M.cattarrhalis
US-09-540-236-646

Query Match 18.0%; Score 353; DB 4; Length 1632;
Best Local Similarity 54.8%; Pred. No. 1,2e-106;
Matches 824; Conservative 0; Mismatches 650; Indels 31; Gaps 5;

Qy	137	CCGGGGCTTATATCCTGATGATGACCTGAACGCCCGGCGTCAACACATTTTGGCT	196
Db	122	CAGCGCGTGAGATCCTTGTTCAAGCATTTGATGAGAGCTGCGAGTATATTTTGGCT	181
Qy	197	ATCCCGGGGGCAATTTTGCCTATGATGAACTGTACCGCTTGAAGCGGGGGG	256
Db	182	ATCAGCGCGGCGCGATTTACATATTTATGATGCTTTATTTAAGCAAGATA-----	232
Qy	257	AAATTGACATTTTGGTGGCCATGATGAACAGAGCTCCCATGGCGGATGGGTAG	316
Db	233	AAATTGACATATTTTGGTGGCCATGATGAACAGCGCGGATATGCGCATGATACA	292
Qy	317	CCAGAGCCACAGGTAAAGTGGAGTTGTTTCCGTACATCTGACACAGGGCGACTAAT	376
Db	293	GTCTGATACACAGGACGAGCGGGCGTGTGTTGGGACATTCAGGTCAAGGTGACCAACA	352
Qy	377	TGGTGAACGGGATTTGCCAATGCCATTTGACTGGGTGCCATGGTGGTATTAATCGAG	436
Db	353	CTGTAAACAGCCATTGCAACTGTCTATATGATTCATATCCAAATGGTGTGTTGGCAGGTC	412
Qy	437	AGGTGGGCGGTGCCATATTGGTAGCGATGCTTTCAGAGAAATTGACATTTTGGCATCA	496
Db	413	AGGTACATCTAGTTTATTTGATGATGATGCGTTTCAAGAGTGTATATGTCGGCGTTT	472
Qy	497	CTTTACCGATGCTTAAAGCATCTCTATGTGTAGTGTAGTGGCGGAGTATGCGTCGATG	556
Db	473	CAGCGCCCATCTGTAAACACAGTTTCAAGTGTCCCATGCGAAGACATTTCCACAAATCA	532
Qy	557	TTACTGAGGCTTTCATCTTGCTAGCACCGGTCGTCGCCGGCTTTGATTCGATATTC	616
Db	533	TCAAAAAGACATTTTATATTTGACAGGCTCGGTCGTCGTCGTCGTCGTCGTCGTCGTC	592
Qy	617	CCAGAGATGTGGGCTTAAAGATGTAGTACATTTCCCTGACCCCGGTGACGTTAATC	676
Db	593	CAAAAGTATGACCAACCCCAAGTATTAATTTGCTTATATCATGCAAAAGACAGTCATTA	652
Qy	677	TACCGGCTTATCGCCCGACGGTTAAAGGTATATCCCGCAAAATTAATGCGCATTCGAT	736
Db	653	TCAATCATATTCAGCTTACTATATTAAGGTCTATGCGGACAAATCAAAAAGCATCGAAA	712
Qy	737	TGTTGAGCAGGCCAGAAATCCCTTGCTTACGTAGGGGAGAGGCGCATTCGCCGCATG	796
Db	713	CGCTACTATAGCCCAACGCCCTGACTGATTCAGGTGGTGGGTGGTGGTGGTGGTGGT	772

797 CCCATGCCAGTGGCAGAAATTTGGGAAAGGTTCCAGTTGCCGGTACACACCCCTGA 856
DB CCACTAAGAGCTGGCTGATTAAGCCATTAAGCTCAACCTGTTACTTAATACCTTA 832
QY 857 TGGGAATTTGGGCTTTTGAAGAAAACATCCCTTTCGGTGGTATGTGGATGATG 916
DB 833 TGGAGCTGGGCTTTATTCAGAGGCTCAGGCGACATTTGTTGGATCTGTGATGATG 892
QY 917 GCCACCGGTATGCCAATTTGGCGCTCAGCGAATGATTTGATTTGAGTGGAGGCCC 976
DB 893 GCACATATGAAGCGAAGACATGACATGATCTGCTGATGATTTTGGCAGTGGGTGC 952
QY 977 GTTTCAGCAGCCGGGTAACTGGCAAACTAGACGATTTGTCAGCCGCGCAAGTATTC 1036
DB 953 GTTTCAGTGAACCGTGTGACCAATATGCAAAAATTTCTGCCCAATGACATCATTC 1012
QY 1037 ACATTTGATGACATCCCGGCGAGGTGGGAAAAACAGGCTCCCGATGTCCTATGTCG 1096
DB 1013 ATATTGATGATGATTCAGCCAGATATCTCAAGAGATTAATGACACATTCGATTCG 1072
QY 1097 GGAATGATGCGCATGTTTGAACAGCTTTTGAACGCGCGCGGAAATGATTAACCCA 1156
DB 1073 GCGATGTAAGCTGTACTAATGAATGCTTGTGCTGATGACAGATTAATGATGTA 1132
QY 1157 CCATTCCTCATACACCCAGGAGGATTAATGTCATGATTCATTCGCGA----- 1207
DB 1133 TCGATGACATGCGCTCATGATTTGGGCAAAATCAATGAATGCGTAAACGCTAG 1192
QY 1208 -----CCGATTACCCCTCCAGGTGCCCACTATGAGATCTATTTGCCCGCAGAG 1261
DB 1193 GTCTTCGTATGATGCAATATGATGATGCTGCAAAACATGCTATCAACCCAGATG 1252
QY 1262 TAGTACAGAAATTTGTCGCAAGC---CCCCATGCTTACTACACACCGATGTGGAC 1318
DB 1253 TGGTGAAGCCCTTATTAAGCTGAGATGTTGAAGATCATTAATCCGATGTTGTC 1312
QY 1319 AAGCCAAATGTCGCGCGCCAGTTT---TGAACATGCGCGCGCGATGATTTCCA 1375
DB 1313 AGCATGATGTTGTCGATGATTTATTAATGATGATGACCGCAATGTTAACT 1372
QY 1376 GTCTGCGTGGATGATGAGTGGCTTTGTTTACCTGCGCCATGGAGCAAGTGGAG 1435
DB 1373 CAGGTGCTTATGATGATGAGTGGCTTTGTTTACCTTATGCTATGAGCAAAATTA 1432
QY 1436 TGGGGGAGAG-CCGTCACTTTCATGATGATGATGCGAGTTCCAAATGAATCTT 1494
DB 1433 ACCCAAAAGACAGTGTGTTGATCACTGCGAAGGCTCATCAATGAATGATCAAG 1492
QY 1495 AACTGGGAACCTTAGCCAGTACACATCCAGTTAAATATTTCTGATTAACGTT 1554
DB 1493 AGCTGTCACTTGTCTAATATATATACAGTTAAATTTTAAATTAATGCTC 1552
QY 1555 GCGAGGGATGTCGCTGATGCGCAAACTTTCTAGAGAAAGCTTATTCCTCTA 1614
DB 1553 AGCTTGGCATGTTAAGCATGCGAGATATGCTTTTGAAGGTGCGCATGCTCAATCT 1612
QY 1615 ACATG 1619
DB 1613 ATATG 1617

RESULT 8
US-09-557-884-1/c
; Sequence 1, Application US/09557884
; Patent No. 6506581
; GENERAL INFORMATION:
; APPLICANT: Fleischmann et al.
; TITLE OF INVENTION: The Nucleotide sequence of
; the Haemophilus influenzae Rd Genome, Fragments
; thereof, and uses thereof
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:

ADDRESS: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: MD
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: 3 1/2 inch diskette
COMPUTER: Dell Pentium
OPERATING SYSTEM: MS DOS v6.22
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/557,884
FILING DATE: 25-Apr-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/476,102
FILING DATE: JUN-5-1995
ATTORNEY/AGENT INFORMATION:
NAME: Michelle S. Marks
REGISTRATION NUMBER: 41,971
REFERENCE/DOCKET NUMBER: FBI86F3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 301-309-8504
TELEFAX: 301-309-8439
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1830121 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-557-884-1
Query Match 16.8%; Score 328.4; DB 4; Length 1830121;
Best Local Similarity 53.1%; Pred. No. 2.3e-96;
Matches 771; Conservative 0; Mismatches 667; Indels 13; Gaps 3;
QY 173 ATGGGGTCAACACATTTTGGCTATCCCGGCGGGCAATTTGGCCATCATGATGAC 232
DB 1652874 AAGCGTATAGATATATTTGGTATCCCGAGCGCGGTATTAATTAATGATGCA- 1652816
QY 233 TGTACCGCTTGAACCGCGCGGGGAAATGACATATTTGTCGCCATGAACAGAG 292
DB 1652815 -----ANTCATGCGCTGGGTGATTTGAACATATTTTGTGCGCACAAAGCCG 1652764
QY 293 CTTCATCCCGCGGATGGGATATGCGAGCGCAGATGAAGTGGAGTTGTTTCGCTA 352
DB 1652763 CGGTGATATGCGGATGGTATGCGACGTTCAACGGGTAAAGTTGATGTGCTAGTA 1652704
QY 353 CATCTGACCAAGGCGCATTAATTTGATGACCGGCAATGCGAATGCCATTTGA 412
DB 1652703 CTTCGCGGCTGTGCAACAAATGCTATCACTGGCATTTTAAACGCTTATACCGATCTG 1652644
QY 413 TCCCATGCTGTGATTAATGAGAGTGGCGGCTCCATGATGCTGATGCTTCC 472
DB 1652643 TCCCATGCTGTATCAATTTGCGGTGAGTATGAGCAATTAATGAGAGTATGCTTCC 1652584
QY 473 AGGAATGACATTTTGGCATCACTTAACGATGTTAAGCACTCTATGTGATCCTA 532
DB 1652583 AAGATGATATGCTGCTATTTCTGTCCTGTTGAACATATGATTAATGCAAAA 1652524
QY 533 GTGCGCGGATATGCTGATGCTTACTGAGGCTTTCACTTCTGCTAGCACCGGTCGT 592
DB 1652523 AGCGCGAGATATTCATCTATCTTTGAAGAAAGCTTTTAAATGATCTACGGGTCTC 1652464
QY 593 CCGGGCGGTTTGAATGATATTTCCAGAGATGTCGCTTGAAGATGATGATGATTC 652
DB 1652463 CAGGTCTGTGTTGTGATATTTCCAAAGACATGTAATCTTAATTTAAATCTCT 1652404
QY 653 CCTGACCCCGGTGATGATCTTACCGGTTATGCCCCAGGTTAAAGTTATCC 712
DB 1652403 ATGAATACCTGATATGCGATGATCTTCTTAATATCAACAGTAATGACCAAG 1652344

713 GACAAATTAATGCGGATTTGCAATTTGGAGAGCCAGAAATCCCTTGTCTACGTAG 772
DB 1652343 GTCAATTAATTAAGCGTTAAAGCACTTTAGTGGCGAAGAAACCAATCTTTTGTGCG 1652284
QY 773 GGGAGGAGGCGATCCCGCCCAATGCCAGGTGACAGAAATTTGCGAAGAGTTCC 832
DB 1652283 GTGGGTGTGCAATCACTGTGATGTAGTGAAGCAATTAATCAAGTTTGAACAAGTTAA 1652224
QY 833 AGTTCCCGGTAAACAACCACTGTATGGGAATTTGGGGCTTTTACGAAACCAATCCCTTT 892
DB 1652223 ATTTACCCGTGATCTCTCATTAATGGGATTTGGGTCTTATCAAGTACGATTAACAAAT 1652164
QY 893 CGGTGGTATGTGGGTATGCAATGGCAACCGCTATGCCACTTTGGCGTACGCAATGTG 952
DB 1652163 TCTTAGGTATGTGGGTATGCAATGTACTTAAAGCAATACGCAATGACAGAAAGTG 1652104
QY 953 ATTCTGTATGTGCAATGGGGGCGCGTTTTCAGACACCGGGTAACTGGCAAACTAGACGAAT 1012
DB 1652103 ATCTTATTTTGGGGAATGGCGCTTGTGTATGATGTAAGCAAAATTAATTAGAAAT 1652044
QY 1013 TTGCTAGCGCGGCAAGTAATTTCAATTCATTCATGATGCAACCGCGCGAGGTGGGAAAAACA 1072
DB 1652043 ATTGCCCCAATGCAAAAGTATTCATTAATGATGATCAACCTCAATTTCTAAAAATG 1651984
QY 1073 GGGCTCCCGATGTGCCATTTGTGGGGATGTACGCGCATGTTTAAACAAGCTTTTGACAGC 1132
DB 1651983 TGCCAGTAGCAATTCATTTGCGGAATGCGAAAAATGTAATGGAAGAAATTTTGGGTT 1651924
QY 1133 GGGCGCGGGAATGGATTAACCCCAACCATCCCATACCAACGAGCATGTTAAATGCA 1192
DB 1651923 TATTGAATGAAGAGGATTAATTAATCTCAACAGATCTTGAAGTGTGGGCAACAAATCA 1651864
QY 1193 TTGATCATTTGGCGGACCGATTAACCCCTCCAGGTGCCCATATGAGGATCTATGAGCC 1252
DB 1651863 ACCATGGAAGAAACAAAAAATTTTGAATTTGACCGCATCTTGGCGTATTAACCGC 1651804
QY 1253 CCCAGGAGTATGATACAGAAATTTGTGCGCAGGCGCGCGATGCTTACACACCGCATG 1312
DB 1651803 AACCAATGCTGGAAGCGGTATATGCGCTCAGAAAGACAGCTTATGTGCTTCTGATG 1651744
QY 1313 TGGGCAACACCAATGTGGGCGGCG---CGAGTTTGTGAACAAATGGCCCCCGCGCATGGA 1369
DB 1651743 TAGGTGACACCAATTTTGTGTGCTTATTCATTTGATGAGAAACCGGTATGGA 1651684
QY 1370 TTTCAGATGCTGTGGGTATGATGGGCTTTGTGTTTACCTGCCCATGGAAGCAAG 1429
DB 1651683 TTAATTTGTGTGTCAGGCACAATGGGTTTGTGTTTTCGCGGCACTTGTGTAAAT 1651624
QY 1430 TGGGAG-TGGGGAGAGAGCGGTCACTTTGATCACTGATGAGATGCCAGCTTCCAAATGATC 1488
DB 1651623 TAGTCACTCTGAAGAACTGTGTGCTTACTGAGGATGATGATTAATCAATGATA 1651564
QY 1489 TTCAAGAACTGGGAACCTTAGCCAGTACGATCCAGTAACTTAATTAATTTCTCAAT 1548
DB 1651563 TTCAAGAGCTTTTACCGCAACACAAATATGSCATTCCTGTTGTCAATTTTGTGTAATA 1651504
QY 1549 ACCGTTGGCAGGGGATGTGTGCTGATGAGCAACAACTTTCTACGAAGAAAGTTATTTCTG 1608
DB 1651503 ATCAATTTCTTGGGAATGTGTAACAAATGAGATTTGATTTATTTCTGCGCATTTGCG 1651444
QY 1609 CTTCATCAATG 1619
DB 1651443 AAACCTATATG 1651433

Owen White
Hamilton O. Smith
J. Craig Venter
TITLE OF INVENTION: The Nucleotide sequence of
the Haemophilus influenzae Rd Genome, Fragments
Thereof, and Uses Thereof
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville,
STATE: MD
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: 3 1/2 inch diskette
COMPUTER: Dell Pentium
OPERATING SYSTEM: MS DOS v6.22
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/643, 990A
FILING DATE: 23-Aug-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/487,429
FILING DATE: 1995-06-07
APPLICATION NUMBER: 08/426,787
FILING DATE: 1995-04-21
ATTORNEY/AGENT INFORMATION:
NAME: Kenley K. Hoover
REGISTRATION NUMBER: 40,302
REFERENCE/DOCKET NUMBER: PB186P1C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 301-610-5790
TELEFAX: 310-309-8439
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1830121 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-643-990A-1

Query Match 16.8%; Score 328.4; DB 4; Length 1830121;
Best Local Similarity 53.1%; Pred. No. 2.3e-96;
Matches 771; Conservative 0; Mismatches 667; Indels 13; Gaps 3;
QY 173 ATGAGGTCMAAACACATTTTGTGCTATCCCGGAGGGCAATTTGCCATCTATGATGAAC 232
DB 1652874 AAGGCGTAGAGATGATTTGTGTATCCCGAGGCGCGGTATTAATTTATGATGCA- 1652816
QY 233 TGTACCGCTTTGAAGGCGGCGGGAATAGACATATTTTGGGGCCATGAAACAAGAG 292
DB 1652815 -----ANTCATAGCGTGGGTGTGATTTGATTTTGTGCGCCAGCAAGACCGG 1652764
QY 293 CTTCCTAGCGGCGGATGGGTAATGCCAGCCACAGTAAGTGGAGTTTGTTCGGTA 352
DB 1652763 CGGTGCAATATGCGCGGATGGTTACGACGTTCAACGCGTAAAGTGGATGTGCTAGTA 1652704
QY 353 CATCTGACACGAGGGGAGCTAATCTGTGTGACCGGAGTTGCCAATGCCATTTGCACTCGG 412
DB 1652703 CTTCGGGCGCTGTGCAAAATNGCTATCACTGCGCATTTTAAACGCTTATACCGATTTCTG 1652644
QY 413 TGCCATGTGTGATTTACTGTGAGAGGTGGCGGTGACATGATGTGATGCAATGCTTTCC 472
DB 1652643 TGCTATGTGTATCACTTGGGTCAGGTTATGAGCAATTAATTTGCGAGTATGCTTTCC 1652584
QY 473 AGAAATTAACATTTTGGCACAACCTTACCGCATGTTAAGACATCTCTATGTGTAGCTGA 532
DB 1652583 AAGAAATGATATGCTGGGATTTTCTGCTCCGTTGTGAAGCATAGCTTTATTTGCAAAA 1652524
QY 533 GTGCGGCGGATATGCTGCAATTTTACTGAGAGCTTTCCATCTTGTCTAGCAACCGTCTGC 592

Db 712 AGGATCCCGCAACCTGCAACCTGAAAGGGCTGGGTGGATGCTGATTATCCCTAT 771
 QY 892 TCGGTGGATATGTTGGATGATGATGCGCAACCCCTATGCAACTTTGCGTCAGCAATGT 951
 Db 772 TATCTGGCATGTGGGATGACAGCGCACTAAGGGCGCCAACTGCGGTACAGAGATGC 831
 QY 952 GATTTCATTGATGCAATGGGGGCCCTTTGCAAGACCGGGTAATCTGGCAACTAAGCA 1011
 Db 832 GATTCGCTGATGCGCGCTGCGCGCGCTTTTGAAGATCGGTAAACCGGTAACTCAATAC 891
 QY 1012 TTTGCTAGCCGCGCCCAATTAATTCATTTGATGATGACCGGCGGAGGTGGGAAAAAC 1071
 Db 892 TTGCGACCGCATGCGCAAGTATCCATGACATGACATGATCCGCGGAGCTGAACAAGCTG 951
 QY 1072 AAGGCTCCCGATGTGCCATTTGTGGGGATGTAACCGCATGTTTAAACAGCTTTTGAG 1131
 Db 952 CCGTACGCGCATATCGGTCTGACAGGCGATCTGAATGCTGCTGCTGCGCGCTGACAG 1011
 QY 1132 CCGGCGCGGGGAATGGAATTAACCCCAATCCCAATCCCAAGGCAATGTTAATATGC 1191
 Db 1012 -----CGTTAGCTATGATGCGCGCGAGCGG 1041
 QY 1192 ATGTATCATTTGGCGGACCGATTACCC-----CTCAGGTGGCCCACTATGAGATCTATT 1248
 Db 1042 AAGTCCGCACTGTGCTCGAGCATGCTGGCGTTAGCATATCCGGTGAAGCATATCAC 1101
 QY 1248 GCGCCCGGAGAGTATGACACGAATTTGTGCGCAGCGCCCGGATGCTTACTACACCA 1308
 Db 1102 GCGCGCTGTCTCTAAGAGAGCTTTCCAGCCGCAAGCCGCGAGAGGTGTGACCA 1161
 QY 1309 GATGTGGGACACACCAAAATGTGGGCGGCCAGTTTAAACATGAGCCCGCCGCGA--- 1365
 Db 1162 GACGTGGGCAACACCAAGATGTGTGCTGCTGACATATGACCTTAACCCCGGAAAAAC 1221
 QY 1366 TGGATTTTCCATGCTGCTGGGTTCGATGAGGCTTTGGTTTAACTGCGCGCATGGAGCC 1425
 Db 1222 TTATACCTCCAGCGGCTTAAGGACCAATGCGCTTGGCGCTGCGCGCGCTGAGCC 1281
 QY 1426 AAGTGGGAGTGGGGGAGCA-GGGTCAATTTGCAATGAGTGAAGTCCAGCTTCCAATG 1484
 Db 1282 CAGGTGGCGCGCCCTTAACGATACGATACGATCTGATCTCCGATGAGGCTCTTCATGATG 1341
 QY 1485 AATCTTCAGAACTGGGAACTCCAGCCAGTACGACATCCAGTAACTAATTAATCTC 1544
 Db 1342 AATGTACAGGAGCTGGGACCGTTAAAGCAGGAGTTAACGCTGAAGTGTGTATCTC 1401
 QY 1545 AATTAAGGTGGCAGGGAGTGTGTGCTCAGTGGCAAACTTTTCTAGAGAAAGCTTAT 1604
 Db 1402 GACAAACAGCGTTAAGGATGTTGACATGAGCAGCGTGTTTTCCAGAAAGTAT 1461
 QY 1605 TCTGCTTCAATATCTCCAGGGCATGCGCAATTAATCTCTGTAAGAGCTTATGCG 1664
 Db 1462 ACGGAACCACTGTAC---TGATTAACCTGATTTCCAGCGTGGCAGCGCTTGGGT 1518
 QY 1665 ATCAAGGGTATTACTGTGCGCAGCGGAAAGATTTGGCCCGCGATGCGGAATGTCTA 1724
 Db 1519 ATTCCTGGCAACACATCAACCGTAAGACAGGTTGAAGGGCACTGACACCATGCTT 1578
 QY 1725 GCCCAATATGCTCTGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 1784
 Db 1579 TCGAACCGAGGGCCATACCTGCTTCAATCTCAATCGACGAATTTGAAGATGTCTGGCG 1638
 QY 1785 ATGATGCGCCCGGCAATGATGATGATGATGATGATGATGATGATGATGATGATG 1821
 Db 1639 TTGTGCGCGCGCGCGCAATGATGATGATGATGATGATGATGATGATGATGATG 1675

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
 ; FILE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: GTC09-03PA
 ; CURRENT APPLICATION NUMBER: US/09/328,352
 ; CURRENT FILING DATE: 1999-06-04
 ; NUMBER OF SEQ ID NOS: 8252
 ; SEQ ID NO 3942
 ; LENGTH: 1731
 ; TYPE: DNA
 ; ORGANISM: Acinetobacter baumannii
 US-09-328-352-3942

Query Match 16.4%; Score 321.6; DB 4; Length 1731;
 Best Local Similarity 53.7%; Pred. No. 3.6e-96;
 Matches 808; Conservative 0; Mismatches 674; Indels 22; Gaps 6;

QY 128 AAGCGCAACCGGGCTTATATCTGATGATGATGATGATGATGATGATGATGATGATG 187
 Db 11 AACTTTTATCTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 70
 QY 188 TTTTGGCTATCCCGCGGGCAATTTTGGCCATATGATGATGATGATGATGATGATGATG 247
 Db 71 TTTTGGTTATCCAGGGCGGAGTATTAATTTATGATGATGATGATGATGATGATGATGAT 130
 QY 248 CCGGCGGGGAAATGACATATTTTGTGCGCCATGACAGAGCTTCCCATGCGCGG 307
 Db 131 -----AAATCAATCATTAACCTGCTGCGCATGACAGCTGCTGATATGACAG 181
 QY 308 ATGGGTATGCAAGGCCAGCAAGTAAAGTGGAGTTTGTTCGGTACATGACAGAGGG 367
 Db 182 ATGCTCTACTGCGGTGCTACAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 241
 QY 368 GCACTAATCTGTGATCGGCGATTTGCCAATGCCAATTTGATGATGATGATGATGATGATG 427
 Db 242 CAACCAATATGACATCAATCAATTTGCAACAGCTTATATGATGATGATGATGATGATG 301
 QY 428 TTAATGAGAGGTGGCGCGGATGATGATGATGATGATGATGATGATGATGATGATGATG 487
 Db 302 TGTGTGCGAGGTGATCAATCAATCTGATGATGATGATGATGATGATGATGATGATG 361
 QY 488 TTGCAATCACTTACCGATGTTAAGACATCTATGATGATGATGATGATGATGATGATG 547
 Db 362 TTGATTTATACGCTCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 421
 QY 548 CTGCAATTTTAAAGAAAGCAATTTTATATGATGATGATGATGATGATGATGATGATG 607
 Db 422 CTGCAATTTTAAAGAAAGCAATTTTATATGATGATGATGATGATGATGATGATGATG 481
 QY 608 TCGATATTTCCAGAGATGTTGGCTTAAAGATGATGATGATGATGATGATGATGATGATG 667
 Db 482 TCGATATTTCTAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 541
 QY 668 ACGTATATCTACCGGGTATGCGCCCAAGCTTAAAGTAAATGATGATGATGATGATGATG 727
 Db 542 AAGTGAAGATGCGTTCAATCAACCACTTACGCGCGCACTGTGTTCAAAATTCGTAAG 601
 QY 728 CATTCATATTTGAGAGAGGAGCAAGATCCCTTGTCTGATGATGATGATGATGATGATG 787
 Db 602 CAATGATGATATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 661
 QY 788 CCGCAATGCGCATGCGCCAGTGCAGAAATTTGCGGAAAGTTCAGTTGCGCGTAACA 847
 Db 662 AAGGGAATGCTTCTGCTTATTAACGAACTTCCGATTTACTGTTATTCGGTAACTA 721
 QY 848 CCAACCGTAAAGGAAATTTGGGGCTTTTGAAGAAACATCCCTTGGGTGGGTATGATG 907
 Db 722 ACHACATTAATGAGGTTAAGGAGATCCCGGCGATGACCAAGTTATATGATGATGATG 781
 QY 908 GTAATGATGCGCAACCGCTATGCGCACTTTGCGCTGACGATGATGATGATGATGATG 967
 Db 782 GTATGACGCTATATGAAAGCCAAATATGAGCAATGACCAATGACGATGATGATGATG 841
 QY 968 TGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1027

Db 842 TTGGGCGCGTTTATGATGACCGCGTAAACAAACCCGCAAAATTCGCCCCGAACTCAA 901
Qy 1028 AAGTAATTCACATTCGACATGCAACCCGCGGAGCGTGGGAAAAAAGAGGCTCCGATTCG 1087
Db 902 AAGTATTCATGTCATATGCAACCCGCGAATTCCTTCAAAACCAATTAATGCGCATATTC 961
Qy 1088 CCATGTGCGGAGATGACGCGCATGTTTGAAGACGTTTTCAGCGCGCGCCGCGAATTCG 1147
Db 962 CGATGCTTGTGTCAGTACGCAAGTTCCTTCAAGAAATGTTGGCAGATTAACATTA 1201
Qy 1148 ATTACCCCA---CCATCCCATACCAACCCGCGATGTTAAATGCGATTCATTCG 1204
Db 1022 ATGTGCTTAAACCAAAATCTGAGACAAATGCGCGCTGTGCTCAATCAATGAATGCG 1081
Qy 1205 GACACGATTTACCCCT---CCAGTCCCGCACTATGAGATACATTTGCCCCGAGAG 1261
Db 1082 GCAAGTTCATGCTTGAATATGAAATTCATCTGATGCACTATGAAGCAGAGAG 1141
Qy 1262 TAGTACAGAAATGCTGCGAGGCCC---GATGCTTACTACACCAACGATTCGAG 1318
Db 1142 TTGTGAAACTTATATATGAAGTGAACCAATGCGCATGTCATTTACTTGAACGTAGTC 1201
Qy 1319 AACACCAATGTTGGCGCGCC---AGTTTGAACCAATGCGCGCGCGCATGATTCGA 1375
Db 1202 AACATCAATGTTGTCATCTTATATATGATCAAAAGTCTCGCAATGATCACT 1261
Qy 1376 GTGCTGCTTGGGATGAGTGGGCTTTGTTTACCTGCGCGCATGCGCAAGTGGAG 1435
Db 1262 CAGGCTGCTTGGACCACTGAGTGTGCTGTTACCTTATGCAATGCGTGGCAAGCTTCCT 1321
Qy 1436 TGGGGAGAGAGG-GTCATTTGCATCAGTGAAGATGCGAGCTTCCAAATGATTCAG 1494
Db 1322 TCCAGACACAAATGTTGTTGATTTAGCGGTGAAGCTTCAATCAAGTGTATTCAG 1381
Qy 1495 AACTGGGAACCTTACCCGATGACGACATCCAGTTTAAACTATTTCTCAATTAACGTT 1554
Db 1382 AACTTCAACCTGTAAACATATGCTTGAATGTAATAAATCTTGCTGTAACAACGCTG 1441
Qy 1555 GCGAGGATGCTGCTGATGAGTGGCAACAACCTTCTACGAAGAACGTTATTCCTCTA 1614
Db 1442 CATTAAGTATGCTGAAACAAATGCGCAAGTATGAACATGAGGTGCTCAATCAAGCTCT 1501
Qy 1615 ACAT 1618
Db 1502 ACGT 1505

RESULT 12
US-10-096-571-11
; Sequence 11, Application US/10096571
; Patent No. 6623944
; GENERAL INFORMATION:
; APPLICANT: RIEPING, MECHTHILD
; TITLE OF INVENTION: Processes for the preparation of D-pantothenic acid and/or salts th
; FILE REFERENCE: 21499
; CURRENT APPLICATION NUMBER: US/10/096,571
; PRIOR FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: DE 10112102.4
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: US 60/304,776
; SOFTWARE: PatentIn version 3.1
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 11
; LENGTH: 2111
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: RBS
; LOCATION: (8)..(12)
; OTHER INFORMATION:
; NAME/KEY: mutation

LOCATION: (8)..(8)
; OTHER INFORMATION: Insertion of the base A at position 8
; NAME/KEY: CDS
; LOCATION: (19)..(1665)
; OTHER INFORMATION:
US-10-096-571-11

Query Match 16.3%; Score 318.8; DB 4; Length 2111;
Best Local Similarity 52.8%; Pred.No.3.5e-95;
Matches 873; Conservative 0; Mismatches 717; Indels 64; Gaps 6;

Qy 1172 CATTGGGTCAACACATTTTGGCTATCCCGCGGGGGAATTTTSCCATCATGATGATA 231
Db 58 CAGGGGTGAACACCGTTTTCGTTATCCGGGTGGCGCAATTAATGCGGTTTACATCA 117
Qy 232 CTGTACCGCTTTGAACGCGCGGGGAAAAATGACATATTTTGTGCGCCATGAACAGAA 291
Db 118 TTGTATGAC-----GCGCGCGTGAAGCACTTGCTGTGCGACATGAGAGGAT 165
Qy 292 GCTTCCCATGCGCGCGGATGAGTATCCAGAGCAAGATTAAGTGGAGTTTGTTCGT 351
Db 166 GCGGCAATGCGGCTATCGGTTATCGGCTGATCCGCAAACTGCGGTATGATCGCC 225
Qy 352 ACATCTGACACAGGGGCGACTAATTGTTGACCGGCAATGCGCAATGCGCATTTGGAATCG 411
Db 226 ACGTCTGCTCGCGGCGCAACCACTGATTAACCGGCGTTGGGAGCGCATGTGATTC 285
Qy 412 GTGCCATGCTGTGATTACTGAGAGGTGGCCGTGCATGATGTGATGCGATTC 471
Db 286 ATCCCTGTTGTCATCACTACCGCTCAAGTGTCCGACCGTTATGCGACGAGCATTT 345
Qy 472 CAGAAATGACATTTTGGCATTCATCACTTACCGATCGATCGTATTAAGCACTCCATGTGGTACGT 531
Db 346 CAGAGATGAGATGCTCTGAGATTCGTGTTAGCTTACCAAGCAAGCTTCTGTGTGAG 405
Qy 532 AGTGGCGCGGATATGAGCTTGTGATCTGAGGCTTTCATCTTGTAGCAACCGTCT 591
Db 406 TCGCTGAAGAGTTGCTCCGCGCATTTATGCTGTAAGATTCGACGTTCCACAGCTCAGTGT 465
Qy 592 CCGCGCGCGCTTTGATGATATTCGCAAGATGCGGCTTAAGCAAGATGAGTACAT 651
Db 466 CCGTGGCTGCTGTGATATCCAAAGATATCCAGTACGAGGAGTACCTGGA 525
Qy 652 CCCCTGACCCCGGATGACGTTATCTAACCGGTTATGCGCCCAACGTTAAAGTATCC 711
Db 526 CCGTGGTTACCAACCGTTGAATAAGATGATTTCCCAATGCG----- 571
Qy 712 CGACAAATTAATGCGGATGCAATTTGTTGAGCAGGCGCAAAATCCCTGTCTACGTA 771
Db 572 ---AAGTTGAGCAAGCGCGCAATGCTGGCAAAAGCGCAAAACCATGCTGTAGCT 627
Qy 772 GCGGAGGCGGATGCGCGCATGCGCATGCGCGAGTGCAGAGATTTGCGGAAAGCTTC 831
Db 628 GGTGTGCGCTGGGATATGCGAGGAGTCTGCTTTACAGAGATTTCTGCTACACA 687
Qy 832 CAGTTGCGGTAACAACCTGATGGAATGAGGCTTTTGAAGAAACATCCCTT 891
Db 688 AAAATGCTGCCACCTGACGCTGAAGAGGCTGGGCGGAGTTGAAGAGATTAATCCGTAC 747
Qy 892 TCGGTGGATATGTTGGGTATGATGCGACACCGCTATGCAACTTTGCGGTGACGATGT 951
Db 748 TATCTGGGATGCTGGGAATGATGCGACCAAAAGCGGAACTTGCGGTGAGAGTGC 807
Qy 952 GATTGTGATGCAAGTGGGCGCGCTTTGAGCAACCGGATTAATGCGCAACTAGACAA 1011
Db 808 GACTTGTGATGCGGTGGGTGCAAGTTTGAATGACCGGATGACCGGAAACTGAACCC 867
Qy 1012 TTGTGAGCGCGGCAAGTATTAATGACATGACATCGAACCGGCGGAGAGTGGGAAAAAC 1071
Db 868 TTGCAACAACAGCGATGTTATCAATGATGATGACCGGCGAATTAACAAAGCTG 927
Qy 1072 AAGGCTCCGATGTCCTATGTGCGGAGATGACCATGTTTGAACAGCTTTGAG 1131

Mon Jul 26 12:13:49 2004

us-09-893-033-6.rn1

Page 13

Dp 928 GGTACGAGCACTGTGGCATTTACAGTATATTMAATGCTGTGTTAACGACATTTACAGAG 987

OY 1132 CGGCCCGGGAAATTGGATTACCACCCATCCCACATPACCCAGGCATGGTTAATGCG 1191

Dp 988 CCGTTAA-----ATAATCATTGACTGGCACCTACACTGC 1020

OY 1192 ATTGATCATTTGCGCGGACCGATTACCCCCCTCCAGGTGGCCCCCATTAAGAAGATACATTTGCC 1251

Dp 1021 GGCGAGCTGGCTGAATGAACAATGCTGGCGTTATCGACACATCCCGGTGAAGCTATCTAGCCG 1080

OY 1252 CCCGAGAGGTAGTACACGAAAATTGGTCGACAGGCCCCCGATGCTCATCTACACACCGAT 1311

Dp 1081 CCATTGTTGTTAAAAACAACCTGCGGATCGTAAACCTGGCGATTGGCTGTGACACACGAT 1140

OY 1312 GTGGGACACACCAAATGTGGGCGGCGCCAG---TTTTGAACATTTGGCCCCCGCCGATGG 1368

Dp 1141 GTGGGGGACGACCAAGATGTGGGCGCGCGACACATCGACACACTGGCCCCGAAAATTTC 1200

OY 1369 ATTTCCAGTGTCGTGGTGGGTACGATGGGCTTTGGTTACCTGCGCCATGGAGCCAAA 1428

Dp 1201 ATTACCTTCACCGGCTTAGGCACCATGGGTTTCGGTTTACCGACGGCGGTTGGCGACAA 1260

OY 1429 GTGGGAGTGGGGGACGA-CCGGTCAATTGCATCAAGTGAAGTGGCAGTTCCAAATGAT 1487

Dp 1261 GTTCGACGACCGAAGATACGTGTGTCTGTATCTCCGGTGAACGGCTCTTTATATATGAT 1320

OY 1488 CTTCAGGAAGTGGNAACCCCTAGCCCAAGTACACATCCAGTAAAACTATTATTCTCAAT 1547

Dp 1321 GTGCNAGAGCTGGGACCGCTAAACGCAAGCATTAACGCTTGAATAATGCTTTACTCGAT 1380

OY 1548 AACGGTTGGCAGGGATGTGCTGCTCAGTGGCAAAACTTCTTACGAAAGACGTTATCT 1607

Dp 1381 AACCACCGGTAGGATGTTTCGACNATGGCAGCAACTGTTTTTCAGGAACGATACAGC 1440

OY 1608 GCTTCTACATGTCCCAAGGGCATGCCAGACATTAATCTCTCTGTGAAGCCTATAGCATC 1667

Dp 1441 GAACCAACCTTAC--TGATTAACCCGATTTCCGATGTTAGCAGAGCGCTTGGGATC 1497

OY 1668 AAGGTATTACTGTGCGCAAGCGGAGAAATTTGGCCCCCGGCGATGCGGAAATCTAGCC 1727

Dp 1498 CCTGGCCAACATCACCCCGTAAAGACCAAGGTTAAGGGGCACTGCACCATCTGAC 1557

OY 1728 CACATAGTCTCTGTGTATGATGATGTGTGCTCAAAAAAGATGA AAACTGTTACCTATG 1787

Dp 1558 AGTATGGGCATACCTGCTTCACTGTCTCATTCACGAACTTGAAAGCTGTGGCCGCTG 1617

OY 1788 ATTGCCCCCGGCATGATGATATGCCCCAAATGCTAG 1821

Dp 1618 GTGCCGCTGGCGCCAGTAAATTCAGAAATGTTGG 1651

```

? RESULT 13
? US-10-096-571-13
? Sequence 13, Application US/10096571
? Patent No. 6623944
? GENERAL INFORMATION:
? APPLICANT: RIEPING, MECHTHILD
? TITLE OF INVENTION: Process for the preparation of D-pantothenic acid and/or salts thereof
? FILE REFERENCE: 211499
? CURRENT APPLICATION NUMBER: US/10/096,571
? CURRENT FILING DATE: 2002-03-14
? PRIOR APPLICATION NUMBER: DE 10112102.4
? PRIOR FILING DATE: 2001-03-14
? PRIOR APPLICATION NUMBER: US 60/304,776
? PRIOR FILING DATE: 2001-07-13
? NUMBER OF SEQ ID NOS: 14
? SOFTWARE: PatentIn version 3.1
? SEQ ID NO 13
? LENGTH: 2111
? TYPE: DNA
? ORGANISM: Escherichia coli
? FEATURE:
? NAME/KEY: RBS

```

```

; LOCATION: (8) .. (12)
; OTHER INFORMATION:
; NAME/KEY: mutation
; LOCATION: (8) .. (8)
; OTHER INFORMATION: Insertion of the base A at position 8
; NAME/KEY: CDS
; LOCATION: (1662) .. (1925)
; OTHER INFORMATION:
US-10-096-571-13

```

Query Match	16.3%	Score 318.8;	DB 4;	Length 2111;
Best Local Similarity	52.8%;	Pred. NO. 3.5e-95;		
Matches 873;	Conservative	0;	Mismatches 717;	Indels 64;
				Gaps 6;

OY	172	TATGGGGTCAACA	CATT	TTT	TGGCTATCCGGGGGGGCA	TTT	GCCATCTGATGAGAA	231
Db	58	CAGGGTGGAA	CACCGTTT	TGGTATCCGGGTGGCCAA	TATATGCGGTTTACGATCA			117
OY	232	CTGTACCGCTT	TGAAGCGCGGGGAAA	ATTGACATATTTT	TGTCGCCCATGAA	CAAGA	291	
Db	118	TTGTATGAC-----	GGCGGGTGGAG	CACTT	GTGTCTCCGACATGAC	AGGGT	165	
OY	292	GCTTCCCATG	CGGGCGGANTGGGTAT	GCCAGAGCCACAGTAA	AGTGGGAGTTT	TGTTGGT	351	
Db	166	GGCGCAATGG	GGGCTATGGGTAT	TGGCCCGTGTACCGGCAAA	CTGGGATGTGATGCGC		225	
OY	352	ACATCTGGA	CCAGGGGGCGACTT	ACTTGGTGA	CCGGCATTTGCCAA	TGCCATTTGGACTCG	411	
Db	226	ACGTCTGTCC	GGCGCGCAACCA	CTGATTAACGGGGCTTGGGACG	CACTGTAGATTCT		285	
OY	412	GTGGCCATGG	GTGTGATTA	CTGAGAGGTGGGGCGGTGC	ATGATTTGTATGAGATGCTTC	471		
Db	286	ATCCTCTGTT	GTGCGATCA	CCGGTCAAGTGTCCGACCGTTTATCGG	ACGGACGCAATT	345		
OY	472	CAGGAAATGA	CAATTTT	TGGCATTACCTTACCGATCGTTA	AGCATCTCTATGTGTAGT	531		
Db	346	CAGGAAAGTGA	TGTCCTGGAGTTGT	CTGTAGCCTGTACCAAG	CAACATTTCTGTGTGAG	405		
OY	532	AGTCCGGCGGA	TATGGTCGCATGTTACT	AGGGCTTTCATCTTGGTAGCA	CGGGTGT	591		
Db	406	TGCTGGAAAG	TGTCCGGCGCATTTATGGCTG	AAAGCATTTGACGTTGGCA	GTGCAAGTGTGAGTGT	465		
OY	592	CCCGGGCCGG	TTTGTGATCGATATT	CCCAAGATGTGGCTT	TGAAAGATGTAGTACAT	651		
Db	466	CTGTGTCCGG	TTCGTGCGATATCTCCAA	AGATATCCAGCTAGCCACGGGATCCTGGAA	525			
OY	652	CCCCTCGAC	CCCCGGGTGACGTTAATCT	CCGGGTATGCCCCACAGGT	TAAAGTAATCTCC	711		
Db	526	CCGTGTCTAC	CAACCGTTGAA	AAAGATGACTTCCCA	CATGGC-----	571		
OY	712	CGACAAATTA	TATGGCGCATTTG	CATTTGTGGAGCAGGCTCAGAA	TCCCTTGCTACGTA	771		
Db	572	----AAGTTG	AGCAAGCGCCCGCATGCTGTG	CAAAACGCAAAACCGATGCTGTAGT	627			
OY	772	GGGGAGAGGG	GGAGATCGCGCATATGCC	ATGCCCCAGGTGACAGAA	TTTGGGAAAGGTG	831		
Db	628	GGTGTGGCG	GTATGGGGCAGGCA	GTTCGTCTTACGAA	TTTCTGCTACCA	687		
OY	832	CAGTTCCGGT	ACAA	CCCTGTGATGGGA	ATTGGGCTTTTGCGAA	ACATGCCCTT	891	
Db	688	AAATGTCTG	CAACCTGACAGCTG	AAAGGGCTGGGCCAGTTG	AGACGATTAATCCGTAC	747		
OY	892	TGCGTGGGTA	TGTGTGGTATGATGAG	CGACCGGCTATCCAA	CTTGTGCGTAGAGATG	951		
Db	748	TATGTGGG	CATGCTGGGAATG	CATGACCAAAAGCGCGAA	CTTCCGGGTGACAGATG	807		
OY	952	GATTTGTGAT	TGACGTGGGGCGCGTTT	TGACGACCGGGTAA	CTGTGCAAACTTAGACGA	1011		
Db	808	GACTTGTGTAT	CGCCGGGTGACGTTT	TGATGACCGGGGTGAC	CCGGCAACTGMA	CACC	867	
OY	1012	TTTGTCTAG	CGCGCGCAAGTAA	TTCACATGTACAT	CAACCGCGGAGGTGGGAA	AAAC	1071	
Db	868	TTTGTCTACCA	CGCTGATGTTTATTC	CAATGGAATATGCA	CTTGTGTAAGAAATTA	CAACTC	927	

1072 AGGGCTCCCGATGTCCTATGTCGGGATGTCAGCCATGTTTGAACAGCTTTTGAG 1131
928 COTAGGACATGTCGATTAACAAGGATTAATATGCTGTGTACAGCATTAACAGAG 987
1132 GGGGCGGGGAATTGATTAACCCACCATCCCATACACCCAGGATGTTAAATGCG 1191
988 CCGTTAA-----ATATCATGTGTCGACGCTACACTGC 1020
1192 ATTGATCATGTCGGACCGATTAACCCCTCCAGGTGCCCCACTATGAGATTAATGCC 1251
1021 GCGAGAGTGGGTATGAACATGCTGGGTAAACACATCCCGGTGACGCTATGTAGCG 1080
1252 CCCCAGAGGTATGACAGAAATTGTCGCCAGGCCCGGATGCTTCTACACCCAGAT 1311
1081 CCAATTGTTTAAACAACATGTCGATGCTAAACCTGCGATTCGTCGACAGCAGAT 1140
1312 GTGGACACACACCAATGTCGGCGCCAGCCAG---TTTTGAACAATGCCCCCGCGATG 1368
1141 GTGGGCGACACAGATGTCGGCGCCGAGACATGACACATGCTCCCGGAAAAATTTC 1200
1369 ATTTCAGTCTGCTGCTGGTATGACATGCGCTTTGCTTACCTGCGCATGAGCCAAA 1428
1201 ATTACTCTCAGCGGCTTAGGACACATGCGTTTCGTTTACGACGCGGCTTGCGCACA 1260
1429 GTGGAGATGCGGGGACGA-GCGGTCAATTGATGATGAGATGTCAGCTTCCAAATGAT 1487
1261 GTCCGACGACCGAAACGATCTGTCTGTATCTCCGCTGACGCTCTTTCATGATGAT 1320
1488 CTTGAGAACTGGGAACTTACCCAGTACGACATCCAGGTTAAACTATTAATTTCAAT 1547
1321 GTGCAAGAGTGGGACCGTAAACGCAAGCAGTTACCGTTGAAAAATCGCTTACTGAT 1380
1548 AACGGTTGGCAGGGAGTGTGCTGACGTGCAACAACTTTCTACGAAAGACGTTATCT 1607
1381 AACCAACGTTAGGATGTTGACAAATGCGACGACTGTTTTCAGAAACGATAGAGC 1440
1608 GCTTCTAATCATGTCGCCAGGCGATCCAGACATTAATCTCTCTGTGAGGCTTATGAGC 1667
1441 GAAACCACTCTTAC---TGATTAACCCGATTTCTCATGTTAGCAGCGCTTCTGAGATC 1497
1668 AAGGTATTAATGTCGCGACGCGGAGATTTGGCCCCCGGATGCGCGAAATGCTAGCC 1727
1498 CCGGCCAACACATCACCCGTTAAAGACAGGTTAAAGCGGACACGACATGCTGATC 1557
1728 CACATGTCCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1787
1558 AGTATGAGGCTATGCTGCTTCACTGCTCAATGACGACGATGAGAACTGCTGAGCGCTG 1617
1788 ATTGCCCCCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1821
1618 GTGCGGCTGCGCGCAGTATTCAGAAATGTTGG 1651

RESULT 14
US-08-452-075-1
Sequence 1, Application US/08452075
Patent No. 5998178
GENERAL INFORMATION:
APPLICANT: HASHIGUCHI, KEN-ICHI
APPLICANT: KISHINO, HIROKO
APPLICANT: TSUTSUMOTO, NOBUHARU
APPLICANT: MATSUI, HIROSHI
TITLE OF INVENTION: L-ISOLEUCINE-PRODUCING BACTERIUM AND
METHOD FOR PREPARING L-ISOLEUCINE THROUGH FERMENTATION
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESS: SPIVAK, MCCLELLAND, MAIER & NEUSTADT
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, FOURTH FLOOR
CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: USA
ZIP: 22202

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/452,075
FILING DATE: 26-MAY-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 116340/1994
FILING DATE: 30-MAY-1994
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F
REGISTRATION NUMBER: 08/452,075
REFERENCE/DOCKET NUMBER: 10-741-0
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 413-3000
TELEFAX: (703) 413-2220
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ. ID NO. 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2841 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 957..1052
FEATURE:
NAME/KEY: CDS
LOCATION: 1195..2838
US-08-452-075-1

Query Match 15.8%; Score 309.2; DB 2; Length 2841;
Best Local Similarity 52.4%; Pred. No. 6.7e-92;
Matches 867; Conservative 0; Mismatches 723; Indels 64; Gaps 6;

172 CATGGGATCAACACATTTTGGCTATCCGCGGGGCAATTTTGGCCATGTATGANGAA 231
1234 CAGGTGTGACACCGCTTTTGGCTATCCGCGTGGCGCAATTTGCGCGTTAGACATCA 1293
232 CTGTAACGCTTTGAAGCGCGGGGAAATTGACATATTTTGGTGGCCATGAAACAGCA 291
1294 TTGTATGAC-----GGCGCGCTGAGACCTTGCTATGCGGACATGAGCAGGGGT 1341
292 GCTTCCCATGCGCGCGATGGGTATGCGCAGACCAAGGTTAAAGTGGAGTTGTTGGGT 351
1342 GCGGCAATGCGCGCTATGCTGTATGCTGTGCTACCGGCAAACTGCGCTATGATGCGC 1401
352 ACATGTGACACAGGGCGCATTAATCTGTGACACCGCATTTGCCAATGCCATTTGGACTCG 411
1402 ACGTCTGGGCGCGGCGCAACCACTGATAACCGGCGCTTGGGAGCGCACTGTGATGCC 1461
412 GTGCCATGTCGTATTAATGAGAGAGTGGGCGCGTGGCATGATTTGATGAGATGCTTTC 471
1462 ATCCCTGTTTGGCATACCGGCTCAAGTGTCCGACCGCTTATGCGGACATGACATTT 1521
472 CAGAAATTTGACATTTTGGCATCACTTACCGATCTGTTAGACATCCATATGCTGATAGT 531
1522 CAGAAAGTGAATGCTCTGGAATGTCGTTACCTGTACCAAGCATAGCTTTCTGCTGACG 1581
532 AGTGGCGGATATGCTGCTGCAATTTTACTGAGGCTTTCATCTTCTGACAGCGGTGT 591
1582 TCGCTGGAAGAGTTCCCGCGCATCAATGCTAAGATTCGACGTTGCTGCTCAGGTGT 1641
592 CCGGCGCGGTTTGTATGATGATTTCCCAAGATGCGGCTTAGAAGAAATGAGATCAAT 651
1642 CTTGATGCGGTTCTGTGATATCCCAAAAGATATCAGTTAGCAGGCGTGAACCTGGA 1701
652 CCGCTGACCCCGTACGTTAATCTACCGGCTTATGCGCCCAAGGTTAAAGATATCCC 711

Db 1702 CCGTGGTTCACCAACCGTTGAAAAAGAGTCTTCCCATGCGG----- 1747
Qy 712 CGACAATTAATGCGGCATTTGCATTTGAGCAGGCGCAGAAATCCCTTGCTCACTGA 771
Db 1748 ----AAGTTGAGCAAGCGCGCCAGATGCTGGCAAAAAGGCAAAAACCATGCTGACCTT 1803
Qy 772 GGGGGAAGGGGCGATGCGCGCCATGCGCATGCGGCAAGTGGCGAATTTGCGGAAAGTTC 831
Db 1804 GCGGATGCGGATGAGTATGCGCGAGGACAGTTCGCGCTTGGCGGAATTTCTGCTGCCACA 1863
Qy 832 CAGTGGCGGTAACACCAACCTGATGGGAATTTGGGCGCTTTTGAAGAAAACATCCCTT 891
Db 1864 AAAATGCTGCTGCCATGTAACGTGAAGAGGCTGGCGGAGTAAGAGCAATTAATCCGTAC 1923
Qy 892 TCGGATGATGTTGGGATATGATGAGCAACCGCTATGCGCACTTTGCGGTGACGCGAATGT 951
Db 1924 TATCTGGGATGCTGGGGATGCAACGCAACAAAGCGCAAACTTTCGCGTGCAGAGAGTGT 1983
Qy 952 GATTGTTGATGATGCGGGGCGCGCTTCCAGCAGACCGGATTAATGCGAAACTGACGAA 1011
Db 1984 GACCTGCTGATGCGGATGCGGCGCAGCTTTTGAATGACCGGATGACCGGCAAACTGAAACCC 2043
Qy 1012 TTTGCTAGCGCGCCCAAAATTAATTCATGATGACATGACCGCGGAGTGGGAAAC 1071
Db 2044 TTGCGGCAACACGCGCAGTGTATTCATATGATGATGACCGGCGAATATGAACAAGCTG 2103
Qy 1072 AGGGCTCCGATGATGCGCATTTGGGGGATGAGCCCATGTTTGAACAGCTTTGCGAG 1131
Db 2104 CGTGAGGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2163
Qy 1132 CGGCGCCGGAATTTGATTTACCCCAACCATCCCATACACCGAGCATGTTAAATGCG 1191
Db 2164 CCGTAAATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2196
Qy 1192 ATTGATCAATGCGGACCGATTAACCCCTCCAGGTGCGCCCATGAGATTAATGCGC 1251
Db 2197 GCGCAGCTGCGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2256
Qy 1252 CCGCAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1311
Db 2257 CCGTGTGTTTAAACAACTGCTGCGATGTAACCTGCGGATGCTGATGATGATGATGATGATGATG 2316
Qy 1312 GTGGAGCAACCAAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1368
Db 2317 GTGGGCGACACCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2376
Qy 1369 ATTTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1428
Db 2377 ATCACCTCCAGCGGTTTAAAGATACATGAGGTTTAAAGGTTTAAAGGTTTAAAGGTTTAAAGG 2436
Qy 1429 GTGGAGTGGGGGAGCA-GCGGTCAATTGATGATGATGATGATGATGATGATGATGATGATGAT 1487
Db 2437 GTGCGGCGACGGAAGATACCTGCTGCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2436
Qy 1488 CTTCAAGAACTGGGAACCTTGAACCGCAGTACGATCAGGTTAAACTATTAATTTCTCAAT 1547
Db 2497 GTGCAAGAGCTGGGACCGTAAACGCAAGAGATTAACCGTTGAAGGTTTAAAGGTTTAAAGG 2556
Qy 1548 AACGGTTGGAGGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1607
Db 2557 AACCAAGGTTAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2616
Qy 1608 GCTTCAATCATGATCCAGGAGATGCGAATTAATCTCTCTGATGAAGCTATGAGATC 1667
Db 2617 GAAACCAACCTTAC--TGATTAACCCCGATTTCTCATGATGAGGCGCTTGGGATC 2673
Qy 1668 AAGGGTATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1727
Db 2674 CATGGCCAAACATACCCGGAAGACAGGTTGAAGCGGACCTGGAACCATGCTGAAC 2733
Qy 1728 CACAAATGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1787
Db 2734 AGTATGAGGCGCATACCTGCTTCAATGATGATGATGATGATGATGATGATGATGATGATGATG 2793

Qy 1788 ATTGCCCCCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1821
Db 2794 GTGCGGCTGCGCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2827

RESULT 15
US-09-231-061-1
Sequence 1, Application US/09231061
Patent No. 6214591
GENERAL INFORMATION:
APPLICANT: TOMITA, FUSAO
APPLICANT: YOKOTA, ATSUSHI
APPLICANT: HASHIGUCHI, KENICHI
APPLICANT: ISHIGOOKA, MASAKO
APPLICANT: KURAHASHI, OSAMU
TITLE OF INVENTION: METHODS FOR PRODUCING L-VALINE AND
TITLE OF INVENTION: L-LEUCINE
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
ADDRESSER: P.C.
STREET: 1755 S. JEFF. DAVIS HIGHWAY, FOURTH FLOOR
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/231,061
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/793,441
FILING DATE: 28-FEB-1997
APPLICATION NUMBER: PCT/JP95/01719
FILING DATE: 30-AUG-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-204856
FILING DATE: 30-AUG-1994
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 10-843-0 PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ. ID NO. 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2841 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHEetical: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: ESCHERICHIA COLI
STRAIN: M162
FEATURE:
NAME/KEY: CDS
LOCATION: 957..1055
OTHER INFORMATION: /note= "IDENTIFICATION METHOD: S"
NAME/KEY: attenuator
LOCATION: 1081..1104
OTHER INFORMATION: /note= "IDENTIFICATION METHOD: S"
NAME/KEY: CDS

```

LOCATION: 1195..2841
OTHER INFORMATION: /note= "IDENTIFICATION METHOD: S"
FEATURE:
NAME/KEY: misc feature
LOCATION: 52..57
OTHER INFORMATION: /note= "CLEAVAGE SITE: SmaI,
OTHER INFORMATION: IDENTIFICATION METHOD: S"
FEATURE:
NAME/KEY: misc feature
LOCATION: 2395..2400
OTHER INFORMATION: /note= "CLEAVAGE SITE: KpnI,
OTHER INFORMATION: IDENTIFICATION METHOD: S"
US-09-231-061-1

Query Match      15.8%; Score 309.2; DB 3; Length 2841;
Best Local Similarity 52.4%; Pred. No. 6.7e-92;
Matches 867; Conservative 0; Mismatches 723; Indels 64; Gaps 6;

QY 172 CATGGGCGTCAACACATTTTGGCGTATCCCGGCGGGGCAATTGGCCATCTATGAGAA 231
DB 1234 CAGGGTGTGAACACCGTTTTCGTTATCCGGGTGCGCATTTATGCGGTTTACGATGCA 1293
QY 232 CTGTACCGCTTTGAAGCGCGGGGAAATTGAGCATATTTGTGCGCCATGAA 291
DB 1294 TTGTATGAC-----GGCGGCGTGAAGCACTTGCTATGCCGACATGAGCAGGT 1341
QY 292 GCTTCCCATGGCGGAGATGGGTATGCCAGAGCCACAGTAAAGTGGAGTTTGTGGT 351
DB 1342 GCGGCAATGGCGGCTATCGTTATGCTGCTACCGGCAAACTGGCGTATGATGCC 1401
QY 352 ACATCTGACACCGGGGCGACTACTGTTGTTGACCGGCAATTGGCAATGCCATTTGSACTCG 411
DB 1402 ACGTGTGTCGGGGGCAACCACTGATTAACGGGGCTGGGAGCGACATGTTAGATTCC 1461
QY 412 GTGCCCATGTGTGTGATTACTGAGAGGTGGCGGTGCATGATGTGTAGCATGCTTTC 471
DB 1462 ATCCCTGTGTGTGCTATCACCGGTCAAGTGTCCGACCGTTTACGGACCTGAAGCAATT 1521
QY 472 CAGGAATTTGACATTTTGGCATCACTTACCGATCGTTAGACCTCTATGATGATGAT 531
DB 1522 CAGGAATGTGATGTCCTGTGATGTCCTGTACCTGTACCAAGCATTTCTGTGTGAG 1581
QY 532 AGTGGCGGATATGTGCTGCAATTTGTAAGGCTTTTCAATCTTGTGACACCGGTGCT 591
DB 1582 TCGCTGGAAGATGTCGCCGATCATGCTGATGACATTCGACGTGCTGTGAGGTGCT 1641
QY 592 CCCGCGCGGTTTGTATGATGATTTCCCAAGGATGCGGCTTAGAAGATGTGATGAT 651
DB 1642 CCGTGTCCGTTCTGTGTATGATCCCAAGATATCCAGTTAGCCAGCGGTGACCTGAAA 1701
QY 652 CCCCTCGACCCCGGTGACGTTATCTACCGGTTATGCGCCACGCTTAAAGTATCC 711
DB 1702 CCGTGTTCACCAACCGTTGAAACGAAAGTGACTTCCCAATGCG-----1747
QY 712 CGACAAATTAATGGGGCATTTGCAATTTGTTGAGAGAGCCAGAAATCCCTTGTCTAGCTA 771
DB 1748 -----AAGTTGAGCAAGCGCCAGATGCTGCAAAAGCGCAAAACCGATGCTGTATGCTT 1803
QY 772 GGGGAGGGGCGATGCGCCGCAATGCCATGCCAGGTGACGATTTGGGAAAGGTTTC 831
DB 1804 GCGCGTGGCGGTGTGATGCGGCAAGCAATTCGCGCTTGTGATTTTCGCTGCACACA 1863
QY 832 CAGTTTCCGCTTAACACCACTCTGATGAGAAATTGGGCTTTTGAAGAAACCATCCCTT 891
DB 1864 AAAATGCTGCGCACTGTATGACCTGAAAGGCTGTGGCGCAGTAGAAGCAGATTTATCCGTAC 1923
QY 892 TCGGTGGGTATGTTGGGTATGCAATGAGCCACCGCTATGCAACTTGTGCGTCAAGGAAATG 951
DB 1924 TATCTGGGCAATGCTGGGATGCAAGCACAAAGCGCAAACTTGTGCGGTGCAAGAGTGT 1983
QY 952 GATTTGTTGATGCAATGCGGCGCTTTCAGACACCGGCTAACTGSCAAACTAGAGAA 1011
DB 1984 GACCTGTGATGCGCGTGGCGCAGTTTGTATGACCGGGTGAACCGGCAAACTGAACACC 2043

```

```

QY 1012 TTGTGATCCGCGCCCAAGTATTCATATGACATGACCCCGGCGAGGTGGGAAAAAAC 1071
DB 2044 TTGCGGCCACAGCCAGTGTATTCATATGATATGACCCCGGCAAGAAATGAACAAAGCTG 2103
QY 1072 AAGGCTCCGATGTGCCATTTGTGGGGATGTGATGCCATGTTTTAGAACACTTTTGAG 1131
DB 2104 CGTCAAGGCAATGTGCAATTCAGAGTATTTAAATGCTGTATCCGCAATTACAGAG 2163
QY 1132 CGGGCCCGGGAATTTGATTTACCCCAACCATCCCATACACCCAGGCAATGTTAAATGC 1191
DB 2164 CGGTAAATCAATGTGATCGGCA-----GCAACACTGC 2196
QY 1192 ATTGATCATTTGGCGGACCGATTTACCCCTCCAGAGTCCCACTATGAGAGATATATGCC 1251
DB 2197 GCGCAGCTGCGTATGATGAACATTTCTGGGGTTACGACATCCCGGTGACGCTATTCAGCG 2256
QY 1252 CCCAGAGGTAGTACAGAAATTTGTGCGCAGGCGCCGATGCTACTACACACCGAT 1311
DB 2257 CCGTTGTTTAAACATGTCGATGATGATGTAACCTGCGGATGTGCTGTACACCAAT 2316
QY 1312 GTGGGACAAACCAATGTGGCGGCGCAG---TTTTGAACATGCGCCCGCGATG 1368
DB 2317 GTGGGGCAGCACAGATGTGGGCTGCGGACACATGCGCCACACTGCGCCGAAATTTTC 2376
QY 1369 ATTTCAAGTGTGGCTGGGTGATGAGTGGGCTTTGTTTACCTGCGCGCATGGAGCCAAA 1428
DB 2377 ATCACTCCAGCGCTTTAGTATGATGATGATGATGATGATGATGATGATGATGAT 2436
QY 1429 GTGGAGATGGGGGAGCA--GCGGTCAATTTGATCACTGATGATGATGATGATGAT 1487
DB 2437 GTGGGCGGACCGAAGATATCCGTGTGTATCTCGGTGACGCGCTTTTATGATGAT 2496
QY 1488 CTTCAAGAACTGGGAACCTTACCCAGTACGACATCCAGTTTAAACTATATTCAT 1547
DB 2497 GTGCAAGAGCTGGGACCGTAAACGCAAGATTTACCTGTGAATATGCTTACTCAT 2556
QY 1548 AACGTTGGCAAGGATGTGTGCTGATGAGCAACAACTTTCTAGAGAAAGTTATCT 1607
DB 2557 AACCAACGTTAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2616
QY 1608 GCTTTAACAATGCCAGGACATGCAACATTAATCTCTGTGGAAGCTTATGAGCATC 1667
DB 2617 GAAACCAACCTTAC---TGATTAACCCGATTTCTCATGTTAGCCAGGCGCTTGGCATC 2673
QY 1668 AAGGATTAATCTGTGCGCAAGGAGGAAATTTGCGCCGCGATGCGCGAATGCTAGCC 1727
DB 2674 CATGCCAACAATACACCGGAAAGACAGATGTGAAGCGCACTGACACCATGCTGAA 2733
QY 1728 CAACAATGCTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1787
DB 2734 ACTGATGGGCAATACCTGCTTCATGTCTCAATGCAAGCAACTGTGAGAAAGCTGTGCGCGCTG 2793
QY 1788 ATTGCCCCCGCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1821
DB 2794 GTGCGGCTGGCGCAATTAATCAGAAATGTTGG 2827

```

Search completed: July 25, 2004, 12:54:47
 Job time : 166.901 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 25, 2004, 07:28:18 ; Search time 943.797 Seconds

(without alignments)
10146.437 Million cell updates/sec

Title: US-09-893-033-6

Perfect score: 1955

Sequence: 1 gccatagagagcccatgcgcg.....cataagcccaaatgcaattc 1959

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3216467 segs, 244149694 residues

Total number of hits satisfying chosen parameters: 6432934

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
6: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
19: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	ID	Description
1	1843.4	94.1	1866	US-10-369-493-26409	Sequence 26409, A
2	847.2	43.2	1815	US-10-369-493-43721	Sequence 43721, A
3	841.8	43.0	1815	US-10-369-493-42620	Sequence 42620, A
4	415	19.2	1728	US-10-282-122A-41727	Sequence 41727, A
5	381.6	19.5	1878	US-10-282-122A-21733	Sequence 21733, A
6	379.8	19.4	1866	US-10-282-122A-25560	Sequence 25560, A
7	379.6	19.4	1722	US-10-282-122A-41305	Sequence 41305, A
8	378.8	19.3	1875	US-10-282-122A-11812	Sequence 11812, A
9	373.8	19.1	1725	US-09-815-242-7947	Sequence 7947, Ap
10	373.8	19.1	1725	US-10-282-122A-30524	Sequence 30524, A
11	368.6	18.8	1815	US-09-815-242-5925	Sequence 5925, Ap
12	368.6	18.8	1815	US-10-282-122A-20222	Sequence 20222, A
13	368.2	18.8	1833	US-10-282-122A-26894	Sequence 26894, A
14	368.2	18.8	96109	US-10-672-787-35	Sequence 35, Appl

15	367.8	18.8	1725	13	US-10-282-122A-30934	Sequence 30934, A
16	366.8	18.7	1854	13	US-10-282-122A-28264	Sequence 28264, A
17	366.8	18.7	1857	13	US-10-282-122A-28639	Sequence 28639, A
18	363	18.5	1596	13	US-10-282-122A-32093	Sequence 32093, A
19	363	18.5	1725	17	US-10-275-026A-29	Sequence 29, Ap
20	363	18.5	1725	17	US-10-275-026A-95	Sequence 95, Appl
21	363	18.5	1728	13	US-10-282-122A-29827	Sequence 29827, A
22	358.6	18.3	1782	16	US-10-369-493-3567	Sequence 3567, A
23	357	18.2	1848	15	US-10-156-761-2720	Sequence 2720, Ap
24	357	18.2	9025608	15	US-10-156-761-1	Sequence 1, Appl
25	355	18.1	1725	13	US-10-282-122A-29151	Sequence 29151, A
26	354	18.1	1722	13	US-10-282-122A-24462	Sequence 24462, A
27	350.8	17.9	6506	16	US-10-398-221-3845	Sequence 3845, Ap
28	349.8	17.9	1761	13	US-10-282-122A-13402	Sequence 13402, A
29	345.8	17.7	1554	13	US-10-282-122A-33627	Sequence 33627, A
30	345	17.6	2286	16	US-10-398-221-1981	Sequence 1981, Ap
31	345	17.6	319630	16	US-10-398-221-7	Sequence 7, Appl
32	345	17.6	3011208	16	US-10-398-221-2058	Sequence 2058, Ap
33	342.2	17.5	1713	13	US-10-282-122A-15103	Sequence 15103, A
34	337.8	17.2	1956	13	US-10-282-122A-17528	Sequence 17528, A
35	337.2	17.2	1761	13	US-10-282-122A-14449	Sequence 14449, A
36	330.2	16.9	1642	13	US-10-282-122A-19815	Sequence 19815, A
37	329.4	16.8	1722	9	US-09-815-242-7160	Sequence 7160, Ap
38	329.4	16.8	1722	13	US-10-282-122A-22365	Sequence 22365, A
39	329.2	16.8	1878	9	US-09-738-626-3457	Sequence 3457, Ap
40	329.2	16.8	3309400	9	US-09-738-626-1	Sequence 1, Appl
41	328.4	16.8	1830121	15	US-10-329-960-1	Sequence 1, Appl
42	328.4	16.6	1644	13	US-10-282-122A-23451	Sequence 23451, A
43	325.4	16.5	1722	13	US-10-282-122A-9061	Sequence 9061, Ap
44	323.4	16.5	1701	13	US-10-282-122A-36338	Sequence 36338, A
45	321.4	16.4	1701	13	US-10-282-122A-36338	Sequence 36338, A

ALIGNMENTS

RESULT 1
US-10-369-493-26409
Sequence 26409, Application US/10369493
Publication No. US20030233675A1
GENERAL INFORMATION:
APPLICANT: Cao, Yongwei
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Goldman, Barry S.
APPLICANT: Chen, Xianfeng
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
FILE REFERENCE: 38-10(52052)B
CURRENT APPLICATION NUMBER: US/10/369,493
CURRENT FILING DATE: 2003-02-28
PRIOR APPLICATION NUMBER: US 60/360,039
PRIOR FILING DATE: 2002-02-21
NUMBER OF SEQ ID NOS: 47374
SEQ ID NO 26409
LENGTH: 1866
TYPE: DNA
ORGANISM: Synechocystis sp.
US-10-369-493-26409
Query Match 94.1%; Score 1843.4; DB 16; Length 1866;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1858; Conservative 0; Mismatches 7; Indels 1; Gaps 1;
QY 73 GGTGTGCATCAATTCCTCAATCCCAAAATTTCTGAAACTGTTCTACCAAGG 132
1 GTGGTTCATCAATTCCTCAATCCCAAAATTTCTGAAACTGTTCTGACCAAGG 60
DB 133 CAAACGGGGCTATATCTCTGATGATAGCTGAACGCGATGGGTCAACGATTTT 192
61 CAAACGGGGCTATATCTCTGATGATAGCTGAACGCGATGGGTCAACGATTTT 120
QY 193 GGTATTCGGGGCGGGGCAATTTTGCCCATCTATGATGAATGATACCGCTTTGAAGCGGGC 252

```

Db      121 GGTATCCCGCGGCAATTTTSCCATCTATGAACTGTACCCCTTGAAGCGGC 180
Qy      253 GGGGAAATGAGCATTTTGTGTGCGCAAGAAAGAGCTTCCATGCGGGGANTGG 312
Db      181 GGGGAAATGAGCATTTTGTGTGCGCAAGAAAGAGCTTCCATGCGGGGANTGG 240
Qy      313 TATGCCAGAGCCAGAGTAAAGTGGAGTTGTTTCGGTACATCTGGACCGAGGCGACT 372
Db      241 TATGCCAGAGCCAGAGTAAAGTGGAGTTGTTTCGGTACATCTGGACCGAGGCGACT 300
Qy      373 AACTTGTGACCGGCATTTGCCAATGCGCATTTGGACTCGCGGCCCATGCGGTGATTA 432
Db      301 AACTTGTGACCGGCATTTGCCAATGCGCATTTGGACTCGCGGCCCATGCGGTGATTA 360
Qy      433 GGAAGGTGGCGCGCTGCATGTTGTAAGCATGCTTTCAGAGAAATGACATTTTGGC 492
Db      361 GGAAGGTGGCGCGCTGCATGTTGTAAGCATGCTTTCAGAGAAATGACATTTTGGC 420
Qy      493 ATCACTTACCGATGCTTGAAGCATCTCTATGTGTGTAAGTGGCGGATATGGCTGCG 552
Db      421 ATCACTTACCGATGCTTGAAGCATCTCTATGTGTGTAAGTGGCGGATATGGCTGCG 480
Qy      553 ATTGTACTGAGGCTTTCATCTTGTAGCAACCGGTGTCGCGGCGGTTTGTGATGAT 612
Db      481 ATTGTACTGAGGCTTTCATCTTGTAGCAACCGGTGTCGCGGCGGTTTGTGATGAT 540
Qy      613 ATTCGCAAGAGATGTGGCTTGAAGAAATGTAGTACATTCCTCCGACCCCGGTGACGTT 672
Db      541 ATTCGCAAGAGATGTGGCTTGAAGAAATGTAGTACATTCCTCCGACCCCGGTGACGTT 600
Qy      673 AATCAACCGGGTTATGCGCCGCAAGGTTAAAGTAACTCCCGACAAATTAAGCGGCA 732
Db      601 AATCAACCGGGTTATGCGCCGCAAGGTTAAAGTAACTCCCGACAAATTAAGCGGCA 660
Qy      733 CAATGTTGAGCAAGGCGCAAAATCCCTTGTCTACGTAGGAGGAGGCGATGCGCGC 792
Db      661 CAATGTTGAGCAAGGCGCAAAATCCCTTGTCTACGTAGGAGGAGGCGATGCGCGC 720
Qy      793 AATGCCCATGCGCGAGTGCAGAAATTTGCGGAAGGTTCCAGTTGCGCGGTAACAACACC 852
Db      721 AATGCCCATGCGCGAGTGCAGAAATTTGCGGAAGGTTCCAGTTGCGCGGTAACAACACC 780
Qy      853 CTGATGGGAATTTGGGGCTTTTGAAGAAACCATCCCTTTCGATGGGTATGTTGGTATG 912
Db      781 CTGATGGGAATTTGGGGCTTTTGAAGAAACCATCCCTTTCGATGGGTATGTTGGTATG 840
Qy      913 CATGCCACCGCTATGCCAATCTTTCGCGTCAAGCAATGATTTTGTGATTCAGTGGG 972
Db      841 CATGCCACCGCTATGCCAATCTTTCGCGTCAAGCAATGATTTTGTGATTCAGTGGG 900
Qy      973 GCCCGTTTGAAGCAAGGCGTAACTAGCAAACTAGAGCAATTTGCTAGCCGCGCAAGTA 1032
Db      901 GCCCGTTTGAAGCAAGGCGTAACTAGCAAACTAGAGCAATTTGCTAGCCGCGCAAGTA 960
Qy      1033 ATTCACTTGAATGACCGCGCGAGGTGGAAAAAACAAGGCTCCCGATGTCCTCAAT 1092
Db      961 ATTCACTTGAATGACCGCGCGAGGTGGAAAAAACAAGGCTCCCGATGTCCTCAAT 1020
Qy      1093 GTGGGGGATGATACCGCATGTTTGAAGCAAGCTTTTGAAGGCGCGCGCAATTTGATTA 1152
Db      1021 GTGGGGGATGATACCGCATGTTTGAAGCAAGCTTTTGAAGGCGCGCGCAATTTGATTA 1080
Qy      1153 CCCACCATCCCATACCAACCGAGCATGTTAAATGCAATGATTCATTTGGCGGACCGAT 1212
Db      1081 CCCACCATCCCATACCAACCGAGCATGTTAAATGCAATGATTCATTTGGCGGACCGAT 1140
Qy      1213 TACCCCTTCAGAGTCCCACTATGAGATATCTATGCCCCCGAGAGGTATGTAACAGAA 1272
Db      1141 TACCCCTTCAGAGTCCCACTATGAGATATCTATGCCCCCGAGAGGTATGTAACAGAA 1200
Qy      1273 ATTGGTCCGACAGGCCCGCATGCTATACACCAACCGATGTTGGACAAACCAATGTGG 1332

```

```

Db      1201 ATTGTGCCAGGCCCCCGATGCTTACTACACCAACCGATGTGGACAAACCAATGTGG 1260
Qy      1333 GGGGCCAGTTTGTGAAAGATGAGCCCGCGCATGATTTTCAGTGTGCTTGGGTAACG 1392
Db      1261 GGGGCCAGTTTGTGAAAGATGAGCCCGCGCATGATTTTCAGTGTGCTTGGGTAACG 1320
Qy      1393 ATGGGCTTTGGTTTACCTGCGCGCATGAGGACCAAAAGTGGAGTGGGGGACGAAAGCGT 1451
Db      1321 ATGGGCTTTGGTTTACCTGCGCGCATGAGGACCAAAAGTGGAGTGGGGGACGAAAGCGT 1380
Qy      1452 ATTGTCATCAGTGGAGATGCCAGCTTCCAAATGATCTTCAAGAACTGGGAACCTTAGCC 1511
Db      1381 ATTGTCATCAGTGGAGATGCCAGCTTCCAAATGATCTTCAAGAACTGGGAACCTTAGCC 1440
Qy      1512 CAGTACGATCAGGTTTAAATCTATTTCTCAATTAACGTTGGCAGAGGAGTGTGCT 1571
Db      1441 CAGTACGATCAGGTTTAAATCTATTTCTCAATTAACGTTGGCAGAGGAGTGTGCT 1500
Qy      1572 CAGTGGCAACAACTTCTTACGAGAAAGTATCTGCTTCTTAACATGTCCAGGGCATG 1631
Db      1501 CAGTGGCAACAACTTCTTACGAGAAAGTATCTGCTTCTTAACATGTCCAGGGCATG 1560
Qy      1632 CCAGACATTAATCTCTCTGTGAAGCTTATGSCATCAAGGATTAATCTGTGCGCAAGCG 1691
Db      1561 CCAGACATTAATCTCTCTGTGAAGCTTATGSCATCAAGGATTAATCTGTGCGCAAGCG 1620
Qy      1692 GAAATTTGGCCCCCGCATGCGCAAAATGCTAGGCCAATGTCCTGTGTGATGAT 1751
Db      1621 GAAATTTGGCCCCCGCATGCGCAAAATGCTAGGCCAATGTCCTGTGTGATGAT 1680
Qy      1752 GTGTGTGTCAAAAAAGATGAATACTGTACCTATGATTTGCCCCCGCATGAGTAATGCC 1811
Db      1681 GTGTGTGTCAAAAAAGATGAATACTGTACCTATGATTTGCCCCCGCATGAGTAATGCC 1740
Qy      1812 CAAATGCTAGGTTTACCGGAATGCGGTACNGACAAATGTTCCCGGATGTTGAGAGTC 1871
Db      1741 CAAATGCTAGGTTTACCGGAATGCGGTACNGACAAATGTTCCCGGATGTTGAGAGTC 1800
Qy      1872 AACCATGCGCAACCAAAATTTTCATCACCCATGTTTCTGTGTTGTGAGCCCAA 1931
Db      1801 AACCATGCGCAACCAAAATTTTCATCACCCATGTTTCTGTGTTGTGAGCCCAA 1860
Qy      1932 CTCTAA 1937
Db      1861 CTCTAA 1866

RESULT 2
US-10-369-493-43721
/ Sequence 43721, Application US/10369493
/ Publication No. US20030233675A1
/ GENERAL INFORMATION:
/ APPLICANT: Cao, Yongwei
/ APPLICANT: Hinkle, Gregory J.
/ APPLICANT: Slater, Steven C.
/ APPLICANT: Goldman, Barry S.
/ APPLICANT: Chen, Xianfeng
/ TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
/ FILE REFERENCE: 38-10(52052)B
/ CURRENT APPLICATION NUMBER: US/10/369,493
/ PRIOR FILING DATE: 2003-02-28
/ PRIOR APPLICATION NUMBER: US 60/360,039
/ NUMBER OF SEQ ID NOS: 43721
/ SEQ ID NO 43721
/ LENGTH: 1815
/ TYPE: DNA
/ ORGANISM: No. US20030233675A1toc punctiforme
US-10-369-493-43721

Query Match 43.2%; Score 847.2; DB 16; Length 1815;
Best Local Similarity 67.5%; Pred. No. 1,3e-273;

```

Matches 1221; Conservative 0; Mismatches 584; Indels 4; Gaps 2;

QY 126 CCACGCGCAACCGGGCTTATATCTGATGATAGCTGAAGCGCATGGGTCACAA 185
Db 9 CAAAGCTGATCTGGCGGTTTGTCTACTTGAACGCTTCAATCCGACGCGGTGATTA 68
QY 186 CATTTTGGCTATCCCGCGGGGCAATTTTGGCCATCTATGATGAATGTACCGTTGA 245
Db 69 TATTTTGGTATCTCGTGGGGCAATTCACCAATTTATGACGACCTGTATAGGTGA 128
QY 246 AGCGGCGGGGAAATGAGCATATTTTGGTGGCCATGAACAAGAGCTTCCATGCGGC 305
Db 129 AGCAATGTGTCTATTAGCATATTTCTGTAGACACGAAACAGGCGACCCATGCGC 188
QY 306 GGAATGGATATGCCAGAGCCACAGTAAAGTGGAGTTTGTTCGGTACATCTGACGAG 365
Db 189 AGACGGTATAGCCCGCTGACATGGGAAGGTAGAGATGCTTTGGCATCTTGGCCAG 248
QY 366 GAGGATATCTGTGACCGGCAATGCCAATGCCATTTGGAATGGTGGTCCATGGTGT 425
Db 249 AGCACTAATCTGTGACAGGATGCGCACCGCTACATGATGATCAATCCCATATGAT 308
QY 426 GATTACTGAGAGGTGGCCGTGCCATGATTTGTAGCGATGCTTCCAGAAATTCAT 485
Db 309 GGTACGGGGCAGGTAGACGTCCTCCATTTGTACAGATGCTTCCAGAAACGATAT 368
QY 486 TTTTGGCATCACTTACCGATGTTAGACATCTCTATGTGTAGTGTAGTGGCGGATAT 545
Db 369 TTACGCAATTAAGCTAACCATAGTAAAGCATCTTATGATGCGCTGACCCCTAAAGAT 428
QY 546 GCGTCGATTTGTTACTGAGGCTTTTCATCTTGTCTAGCACCGGTGCTCCGCGCGT 605
Db 429 GCGCGAATTTGTCGCGAAGCTTTTCAATGCTGACATGCGGCGACCAAGGCGCAT 488
QY 606 GATCATATTTCCCAAGATGTGCGCTTAGAAGATGTGATCATTTCCCTGACCCCG 665
Db 489 GATGATGTTCCTCAAGATGTGATGCTTTAGAAATTTGATCTATGACCTGTAACGAG 548
QY 666 TGAAGCTATCTAACCGGGTTATCGCCCAACCGTTAAAGTATCCCGCAAAATTAAGC 725
Db 549 TTCAATGATGTTACGGGTTATCGCCCAACCGTTAAAGGAAATCCCGCAAAATTAAGC 608
QY 726 GGCATATGCAATTTGTGACAGCGCCAGAAATCCTTGTCTACATGAGGAGGAGCGAT 785
Db 609 TGCATATGATGATTAATGAAAGTGGCGCTCTGATGATGATGATGATGATGAT 668
QY 786 GCGCCCAATGCGCATGCGCAAGTGCAGAAATTTGCGGAAAGTTTCCAGTTGCGGATAC 845
Db 669 GCGACCTGTGCTCAAGAAATTAAGAACTACCTGATTAATCAATCCCTGTAC 728
QY 846 AACCAACCTGATGAGAAATTTGGGCTTTTGAACGAAACATCCCTTGGTGGGATGTT 905
Db 729 CACCACTTAATGAGGATGCGTCTTGAAGAAATCAATCCCTGATGATGATGAT 788
QY 906 GGGTATGATGCGCACCGCTATGCAACTTTGCGCTGACGAAATGATTTGTTGATTC 965
Db 789 GGGAGTGAACGCGCATGCGTACCTTACTGATGATGATGATGATGATGATGAT 848
QY 966 AGTGGGGGCGGCTTTGACGACCGGGTAACTGCGAAATGCAATTTGCTACGCGGC 1025
Db 849 CGTGGCGCTGATTTGACGATGCGCTTACAGCAAGTTAGCAATTTGCGCTCCCGGC 908
QY 1026 CAAATTAATTCATTTGATGATGACCGCGGAGGTGGAAAAAACAAGGCTCCCGATG 1085
Db 909 CAAATTAATTCATTCATGATGATGACCGCAAGTGGCAAAACCGCATTTCCGAGT 968
QY 1086 GCGCATTTGGGGGATGACCGCATGTTTGAAGAGTTTGAAGCGGGCGCGGAGAT 1145
Db 969 GCGCATCGTGGCGATGCGGAAAGTGTGATGATGATGATGATGATGATGATGAT 1028
QY 1146 GATTAACCCACCATTCACCATACCAAGGATGTTAATGATGATGATGATGATGAT 1205
Db 1029 AGCGCTTAAGCTACCCCTAATCAAAACCAAGAAATGTTAATCTTTAATGCTTGGC 1088

QY 1206 GACCGATTACCCCTCCAGGTCGCCCACTATGAGGATTAATTTGCCCCCAGGAGGAT 1265
Db 1089 CGATGATATCTTAATATGTCACAGCACCCGACGATTTCAACCAAGAGTAT 1148
QY 1266 ACAGAAATTTGTGCGAGGCGCCCGATGCTACTACACCAACGATGAGAACCA 1325
Db 1149 TGTAGAGTCAATGCGCAACGACCAACGCTTCTACACCAAGATGTTGAGCAATCA 1208
QY 1326 AATGTGGCGCGCCAGTTTGAACATATGCCCCCGCATGATTTCCAGTGTGCT 1385
Db 1209 AATGTGGCGCAAAATTTCTCAAAACGCGCCCAAGCGCTGATTTCTGCGCTGTT 1268
QY 1386 GGGTACGATGGGCTTTGTGTTTACCTGCGCATGAGGACCAAGTGGAGTGGGGAGC 1444
Db 1269 GGGAAAGATGGGTTTGTGCTTACCTGCGCAATGGGCGAAAGTGGCATTTCTGATGA 1328
QY 1445 AGCGTCAATTTGATCATGCTGAGATGCGCATCTTCAATGATATCTTCAAGTGGAAC 1504
Db 1329 AGAGTCAATTTGATGATGCGGTATGCGCATTTTCAATGATGTTGCGAGAACCTTGGTAC 1388
QY 1505 CCTAGCCAGATGACATCCAGGTTAAATCTATATCTCAATPACGTTGGAGAGGAT 1564
Db 1389 ACTTGACAGTATGATTAATATGTCAAGCAATATATATCAATPACGCTGGCAAGGAT 1448
QY 1565 GGTGCTGCTGTCGCAACAACTTTCTACGAAAGCTTATCTGCTTCAACATGTC 1624
Db 1449 GGTGCGCAGTGGCAACAACTTTCTATGCTGAGCTTACTGCTCAACATGGAAT 1508
QY 1625 GGGCATGCGAGATTAATATCTCTGTGAGAGCTTATGAGATCAAGGATTTACTGCGC 1684
Db 1509 AGGATGCGACAGTGAAGTATTTGCAAAAGCTTATGAGATTAAGGATATGATTAAG 1568
QY 1685 CAAGCGGAAATTTTGGCCCGCGCATGCGCGAAATGCTAGCCCAATATGTCCTGTGT 1744
Db 1569 CGATGCGATCAATTAAGATGATGATGCGCAAAATGCTGCGACACAAAGGCGCATGAT 1628
QY 1745 GATGATGCTGTGCTCAAAAGATTAATCTGTTACCTATGATTTGCCCCCGCATGAG 1804
Db 1629 CTTAAATGTCACGTCACGAGACGAAACCTCTATCCAAATGATGCTCCAGCAAGAG 1688
QY 1805 TAAATGCCAATCTGATGTTTACCGAAGTGGCGGATACAGCAATGTCCTCCGATGTT 1864
Db 1689 CAAGCTCAATATGTCGCTGCGTGAAGCAACG---CCAAAGCAGACAGACCACT 1745
QY 1865 GAGTGCACCATTTGCCAAACCAAAATTCATACCCATGCTTCTGTTCTGTGTGTG 1924
Db 1746 TTAATGATGATGATGCAATGCGCAAAATGCTCTTACCAACATCTGCTGATGATG 1805
QY 1925 AGCCAACT 1933
Db 1806 GACGAGCT 1814

RESULT 3
US-10-369-493-42620
; Sequence 42620, Application US/10369493
; Publication No. US20030233675A1
GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; NUMBER OF SEQ. ID NOS: 47374
; SEQ. ID NO 42620

LENGTH: 1815
TYPE: DNA
ORGANISM: Anabaena PCC7120
FEATURE:
NAME/KEY: unsure
LOCATION: (1)..(1815)
OTHER INFORMATION: unsure at all n locations
US-10-369-493-42620

Query Match 43.0%; Score 841.8; DB 16; Length 1815;
Best Local Similarity 67.3%; Pred. No. 8,66-272;
Matches 1215; Conservative 0; Mismatches 586; Indels 4; Gaps 2;

QY 127 CAACGGCAACCCGGGCTTATATCTGATGATGCTGAAACGCCATGGGCTCAACAC 186
DB 10 CAACGGGTAAGTGGGCTTGGCGCTTCTGATGATGCTGCGCCACGGCGTTGAGTAT 69
QY 187 ATTTTGGCTATCCCGGCGGGGCAATTTTGGCCATCTATGATGATGATGATGATGAT 246
DB 70 ATCTTTGGTATCTCTGGTGGGGCGATTTCTGCCGATTTATGATGATGATGATGATGAT 129
QY 247 GCGGCGGGGGAATTTGAGCATATTTTGGTGGCCATGAAACAAGAGCTTCCCATGGCG 306
DB 130 GCAACTGGCAGCCTTAAACATATTTCTAGTGAACACACACACAGAGAGCTCATGCTGCT 189
QY 307 GATGGGTATGCGCAGAGCCACAGGTAAAGTGGAGTTTGTTCGGTACATGATGACAGG 386
DB 190 GATGGGTATGCTCGGCAACCGGAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 249
QY 367 GCGACTAATCTTGATGACCGGCGATTTGCCAATGCCCATTTGGAATGCGTGGCCCATG 426
DB 250 GCAACTAATCTTGATGACCGGCGATTTGCCAATGCCCATTTGGAATGCGTGGCCCATG 309
QY 427 ATTATGAGAGAGTGGGGCGGTCGCAATGATTTGATGATGATGATGATGATGATGATGAT 486
DB 310 GTTACTGGGCAAGTACCCCGGAAATGATGATGATGATGATGATGATGATGATGATGAT 369
QY 487 TTGGCATCACCCTTACCGCATGCTTAAAGCATCTTATGATGATGATGATGATGATGATGAT 546
DB 370 TACGGTATTTACTCTACCATTTGTAAGCATCTCTATGATGATGATGATGATGATGATGAT 429
QY 547 GCTGGCATTTGTTACTGAGGCTTTTCATTTCTGATGACACCGGTGCTCCCGGGCGGTTTG 606
DB 430 GCGGCAATTTGGCGGAAAGCATTTCCACATGACGACACAGGCGCTCCAGGGCGAGTTTG 489
QY 607 ATCGATTTCCCAAGAGATGTTGGGCTTGAAGATGATGATGATGATGATGATGATGATGAT 666
DB 490 ATTGATGGCGGAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 549
QY 667 GACGTTAATCTACCGGGTATTCGCCCAACGGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 726
DB 550 TCAATTAATTAACGGGATATGCGCCCGGCGTGAAGGAAATCCACGCAATTAATTAACG 609
QY 727 GCATTGCATTTGTTGAGACAGGCGCAAAATCCCTTGCCTACGTAAGGGGGAAGGGGCGATC 786
DB 610 GCGATTGAATGATCGCGAAAGCGCGTCTTACATAGCTTGGTGGTGGGGGCGATC 669
QY 787 GCCGCAATGCCCATGCGCAGGTGACAGAAATTTGGGAAAGGTTCCAGTTGCGGTAACA 846
DB 670 GCATGCGCGCTCATGACAGAAATTAAGAACTGCGCGAGTATTAATTAATCCCGTACT 729
QY 847 ACCACCTCATGAGGAAATGGGGCTTTTACGAAACCATCCCTTGGGTTGGGATGATG 906
DB 730 ACCACCTCATGAGGAAATGGGGCTTTTATGAACCATCCCTTATCCCTTGGGATGATG 789
QY 907 GGTATGCAATGGCCACCGCTATGCCAATTTGCCGTCAGAGATGATGATGATGATGATGATG 966
DB 790 GGGATGCAACGACAGCTTACCTTACCTTGGTGAATGATGATGATGATGATGATGATGATG 849
QY 967 GTGGGGGCGGCTTTGACAGACCGGATTAATGCGAAACTAGACGATTTGCTAGCGCGGCC 1026
DB 850 GTGGGGTGTAGTGTGATGACCGTGTACTGCGCAANTTAAATGATGATGATGATGATGATG 909

QY 1027 AAAGTAAATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1086
DB 910 AAAGTAAATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 969
QY 1087 CCCATTGTTGGGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1146
DB 970 CCCATTGTTGGGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1029
QY 1147 GATTAACCCCAATCCCAATCCCAATCCCAATCCCAATCCCAATCCCAATCCCAATCCCA 1206
DB 1030 AATGCCAAATCAATCCCAATCCCAATCCCAATCCCAATCCCAATCCCAATCCCAATCCCA 1089
QY 1207 ACCGATTAACCCCAATCCCAATCCCAATCCCAATCCCAATCCCAATCCCAATCCCAATCCCA 1266
DB 1090 GAGGATTAACCCCAATCCCAATCCCAATCCCAATCCCAATCCCAATCCCAATCCCAATCCCA 1149
QY 1267 CACGAAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1326
DB 1150 GTGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1209
QY 1327 ATGTTGGGCGGCGGCTTTTGAACAATGCGCGCGCGGATGATGATGATGATGATGATGATG 1386
DB 1210 ATGTTGGGCGGCGGCTTTTGAACAATGCGCGCGCGGATGATGATGATGATGATGATGATG 1269
QY 1387 GGTACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1446
DB 1270 GGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1329
QY 1447 C-GGTATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1505
DB 1330 CAAGTAACTGATTAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1389
QY 1506 CTAGCCAGTAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1565
DB 1390 GGTGACATATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTA 1449
QY 1566 GTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1625
DB 1450 GTAGCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1509
QY 1626 GGCATGCGCAGACATTAATGCTGTTGATGATGATGATGATGATGATGATGATGATGATG 1685
DB 1510 GGGATGCGCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1569
QY 1686 AAGCGGGAATTTGGGCGGCGGATGCGGGAATGCTAAGCCCATGATGATGATGATGATGATG 1745
DB 1570 CACAGAGACATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTA 1629
QY 1746 ATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1805
DB 1630 TTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1689
QY 1806 AATGCCCAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1865
DB 1690 AATGCCCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1746
QY 1866 GAGTGAACATTTGCCAAACCCAAATTTGATGATGATGATGATGATGATGATGATGATG 1925
DB 1747 TATTTGAGCATTTGTTGGGCGGCTTAAATTAACCAATTTGATGATGATGATGATGATG 1806
QY 1926 GCCAA 1930
DB 1807 ACTTA 1811

RESULT 4
US-10-282-122A-41727
; Sequence 41727, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl

/ APPLICANT: Haselbeck, Robert
 / APPLICANT: Ohlsen, Kari
 / APPLICANT: Zyskind, Judith
 / APPLICANT: Wall, Daniel
 / APPLICANT: Trawick, John
 / APPLICANT: Carr, Grant
 / APPLICANT: Yamamoto, Robert
 / APPLICANT: Forsyth, R.
 / APPLICANT: Xu, H.
 / TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
 / FILE REFERENCE: ELITRA 0344
 / CURRENT APPLICATION NUMBER: US/10/282,122A
 / CURRENT FILING DATE: 2003-02-20
 / PRIOR APPLICATION NUMBER: 60/191,078
 / PRIOR FILING DATE: 2000-03-21
 / PRIOR APPLICATION NUMBER: 60/206,848
 / PRIOR FILING DATE: 2000-05-23
 / PRIOR APPLICATION NUMBER: 60/207,727
 / PRIOR FILING DATE: 2000-05-26
 / PRIOR APPLICATION NUMBER: 60/230,335
 / PRIOR FILING DATE: 2000-09-06
 / PRIOR APPLICATION NUMBER: 60/230,347
 / PRIOR FILING DATE: 2000-09-09
 / PRIOR APPLICATION NUMBER: 60/242,578
 / PRIOR FILING DATE: 2000-10-23
 / PRIOR APPLICATION NUMBER: 60/253,625
 / PRIOR FILING DATE: 2000-11-27
 / PRIOR APPLICATION NUMBER: 60/257,931
 / PRIOR FILING DATE: 2000-12-22
 / PRIOR APPLICATION NUMBER: 60/267,636
 / PRIOR FILING DATE: 2001-02-09
 / PRIOR APPLICATION NUMBER: 60/269,308
 / PRIOR FILING DATE: 2001-02-16
 / Remaining Prior Application data removed - See file Wrapper or PALM.
 / NUMBER OF SEQ ID NOS: 78614
 / SOFTWARE: Patent in version 3.1
 / SEQ ID NO 41727
 / LENGTH: 1728
 / TYPE: DNA
 / ORGANISM: Versinia pectis
 US-10-282-122A-41727

Query Match 21.2%; Score 415; DB 13; Length 1728;
 Best Local Similarity 55.8%; Pred. No. 4.7e-128;
 Matches 929; Conservative 0; Mismatches 705; Indels 31; Gaps 6;
 172 CATGGGGTCAACAACATTTTGGCTATCCCGCGGGGCAATTTGGCCATCTATGATGAA 231
 49 CAGGCGGTAAAGCATGTATTCGGTATCCCGCGGGGCGTACTTGATATTACGATGCC 108
 232 CTGTACCGCTTGAAGCGGCGGGGAAATTGAGCATATTTTGGTCCGCATGAAACAAGA 291
 109 CTGCAC-----ACGGTGGAGGCAATCGATCACTGCTGGTGGCAGTAAAGAGT 159
 232 GCTTCCCATGCGGCGATGGTANGCCAGAGCCACAGGTAAAGTGGAGTTTTCGT 351
 160 GGGGTTCACATGGCCGATGTATGCGGGGCAACGGGCGAGGTGGTGTACTGGTTC 219
 332 ACATCTGACACAGGGGCGCACTAATTGTTGACCCGCATTTGCCAATGCCATTTGACTCG 411
 220 AGCTTTGCTCTGGTGAACCAATGCGATACCGGTATGCGACCGCTATATGATTA 279
 412 GTGCCATGTGTGATTAAGTGAAGGTGGCCGTGCCATGATTTGATGAGTCTTC 471
 280 GTGCAATGTGTGCTTTCAGGCCAAGTGCAGAGTCATGATTTGATTTGATGCTTC 339
 472 CAGGAATTTGATTTTGGCATCACTTACCGATCTTAAAGCACTCTTATGTGTACT 531
 340 CAAAGATGTGACATGTGGGATCTCCCGCCAGTGTCAACAAATAGCTTCTTGTAAAG 399
 532 AGTGGCGGATATGGCTGCATTTGTACTGAGGCTTCCATCTTGTAGACCGGATGT 591
 400 CGCAAGAAGACATCCGATGTGTGTGAAAAAGCTTTTATCTGGCATCCACTGGCGCT 459

592 CCCGCGCGGTTTGTATGATATATCCCAAGATGTGGGCTTAGAAGATGTGATCAT 651
 460 CCGGTCGGTGTATGATGACTTCCCAAAAGCACTTTCGGCCCAAGATTAATGCT 519
 652 CCCCTGACCCCGGTGACGTAAATCTACCGGGTATTCGCCCAAGTTAAAGTAAATCC 711
 520 TACGCTATCCAGAGCAATGATGATGCTCTCTATATCTATACGTACGGTGAAGGCACT 579
 712 CGAATAATTAATGGGATTTGCAATTTGTGAGACAGGCCGAATCCCTTGTACTGTA 771
 580 GGGCAATTAAGCGGTGACCTGCAAAAGATTAAGCGGCTTAAAGCGCCATTAATGATGTC 639
 772 GGGGAGGGGCGATTCGCGCCAAATGCGCATGCCAGGTGAGAAATTTGCGAAAGTTC 831
 640 GGTGTGTGCGATTAAGCGCGCTTGCATGAAGATTAATCTGATTTTGGAGAAAGTCT 699
 832 CAGTTGCCGTAAACACACCTTGAATGAGATTTGGGCTTTTGAAGAAACATCCCTT 891
 700 AACTTACCGGTAAACAGCTCTGATGGGCTTAGGCAATTTCCCTGTACGATCCGCAA 759
 892 TCGGTGGTATGTTGGATGACAGGCAACGCTATGCCAATTTGGCCGTACAGGATGT 951
 760 AGTGTGCGATGTTGGGATGCAATGTAATCTTGAAGCCATATGGCCATGCAATACT 819
 952 GATTTGTTGATGCAATGGGGGCGCTTTCAGACACCGGTTAACTGGCAAACTAGACGA 1011
 820 GATCTGATTTTTCGCGGTGTCCTTTTGAAGATGATGATCCACATATTTAGCAAA 879
 1012 TTTGTCAGCGCCCAAGATTAATTCATTTGATGATGATGATGATGATGATGATGATG 1071
 880 TATTCCTGATGCAAGCTGCAATGATGATGATGATGATGATGATGATGATGATGATG 939
 1072 AGGCTCTCCGATGTCATTTGAGGGGATGATGATGATGATGATGATGATGATGATG 1131
 940 GTACCGCGGATATCCGATTTGGGATGATGATGATGATGATGATGATGATGATGATG 999
 1132 CCGGCGCGGATTTGATTAACCCATCCCATCCCATCCCATCCCATCCCATCCCATCC 1191
 1000 TTGTGTGCGAGGTGCGCTGTACCAAGATTTGATGATGATGATGATGATGATGATG 1059
 1192 ATTGATCATGAGGCGACCGGATTAACCCCTCCAGGTCCCGCATATGATGATGATG 1251
 1060 ATTGAACAATGGGCGACCCGATTAATGCTTGGGATTAACAACAAGAGTGGAAATCAAG 1119
 1252 CCCAGAGGTATGATACGAATTTGATGCGCAGGCCCC--CGATGCTTACTACACACC 1308
 1120 CCACAAGCGGTGATTAAGCTTACATGCTGATTAAGGCGATGCTGATGATGATGATG 1179
 1309 GATGTGGGCAACACCAAAATGTTGGGCGGCC--AATTTTGAACAATGGCCCCCGCA 1365
 1180 GATGTGGGCAACACCAAAATGTTGGGCGGCCCTTTTACTATGCTTGTATTAAGCACTG 1239
 1366 TGGATTTCAAGTCTGCTTGGGATGATGATGATGATGATGATGATGATGATGATGATG 1425
 1240 TGGATTAATCTCCGCTGCTTGGGCAAGATGATGATGATGATGATGATGATGATGATG 1299
 1426 AAGTGTGAGTGGGGAAGC-AGGCTGATTTGATGATGATGATGATGATGATGATGATG 1484
 1300 AATTTGCTTTGCTATATGAACCGGTGCTTGTGATGATGATGATGATGATGATGATG 1359
 1485 AATCTTCAAGAACTGGGAACCTTAGCCAGTACGATCCAGGTAAATTAATTTATCTC 1544
 1360 AATATTCAAGAGTATTAACCGCATGCAATTAATTTCCAGTGTGATTAATTAACCTG 1419
 1545 AATAACGTTGGAGGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1604
 1420 AATAACCGCTATGATGAGATGATGATGATGATGATGATGATGATGATGATGATG 1479
 1605 TCTGCTTCAACATGTCACAGGCGATGCGACATTAATCTCCTGTAAGCTATGATG 1664
 1480 TCTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1536

QY 1663 ATCAAGGATATTACTGTGCGCAGGCGGAAGATTGTGGCCCGGCATGCGCGAATGCTA 1724

Db 1517 CATATTGTTGTGGCTATTATATACGGCGGATGAGCGTGAAGTAACTGTGCGACGCACTG

QY 1725 GCCCA-----CAATGTCCTGTGGATGATGTGTGTGTCGTCAAAAAGATGAA 1772

Db 1597 ACACATTATCCGAACCAATCGTTGTGTTCTTTGTGATGACTGTGATGAAACAGAA 1656

QY 1773 AACTGTACCTATGATGTGCCCGCGCATGAGTAATGCCCAATG 1817

Db 1657 CATGTTTACCCGATGCAAGATTCGTGTGTGTGCGCATGACGGAATG 1701

RESULT 5

```

US-10-282-122A-27733
/ Sequence 27733, Application US/10282122A
/ Publication No. US20040029129A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, Hiansu
/ APPLICANT: Zamudio, Carlos
/ APPLICANT: Malone, Cheryl
/ APPLICANT: Haselbeck, Robert
/ APPLICANT: Ohlsen, Karl
/ APPLICANT: Zyskind, Judith
/ APPLICANT: Wall, Daniel
/ APPLICANT: Trawick, John
/ APPLICANT: Carr, Grant
/ APPLICANT: Yamamoto, Robert
/ APPLICANT: Forsyth, R.
/ APPLICANT: Xu, H.
/ TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
/ FILE REFERENCE: ELTRA.034A
/ CURRENT APPLICATION NUMBER: US/10/282,122A
/ CURRENT FILING DATE: 2003-02-20
/ PRIOR APPLICATION NUMBER: 60/191,078
/ PRIOR FILING DATE: 2000-03-21
/ PRIOR APPLICATION NUMBER: 60/206,848
/ PRIOR FILING DATE: 2000-05-23
/ PRIOR APPLICATION NUMBER: 60/207,727
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: 60/230,325
/ PRIOR FILING DATE: 2000-09-06
/ PRIOR APPLICATION NUMBER: 60/230,347
/ PRIOR FILING DATE: 2000-09-09
/ PRIOR APPLICATION NUMBER: 60/242,578
/ PRIOR FILING DATE: 2000-10-23
/ PRIOR APPLICATION NUMBER: 60/253,625
/ PRIOR FILING DATE: 2000-11-27
/ PRIOR APPLICATION NUMBER: 60/257,931
/ PRIOR FILING DATE: 2000-12-22
/ PRIOR APPLICATION NUMBER: 60/267,536
/ PRIOR FILING DATE: 2001-02-09
/ PRIOR APPLICATION NUMBER: 60/269,308
/ PRIOR FILING DATE: 2001-02-16
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 78614
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 27733
/ LENGTH: 1878
/ TYPE: DNA
/ ORGANISM: Mycobacterium leprae
US-10-282-122A-27733

```

Query Match	19.5%;	Score 381.6;	DB 13;	Length 1878;
Best Local Similarity	54.3%;	Pred. No. 8.8e-117;		
Matches 941;	Conservative 0;	Mismatches 759;	Indels 34;	Gaps 7

QY 109 CTGAAAACGTGTTCTCTGACCAACGGCAACCCGGGGCTTATTCGTGATGATAGCCTGAAA 168
Db 109 CTAGAACGCCCTTTCACCTGAGCAACTACCGGCGCTCATCGGTGATTCGGTTCGTGGAA 168
QY 169 CGCATGGGCTCAACCATTTTGGCTATCCCGCGGGGCATTTTCCCATCTATGAT 228

Db	169	GAATCGATGTCGAGGTCATCTTCGGAAATTC	CCGCGCGTGCCTTCCTGATATAGAC	228
Qy	229	GAATCTACCGCTTTGAAAGCGCGCGGGGAAATGAGCATATTTGGTCCCATGAACA		288
Db	229	CCGCTGTCTCTACTCAAAA-----AAGTACGATGATCTCCGGTCCCGACAGACAG		279
Qy	289	GGACCTTCCCATGCGGGCGGATGGGTATGGCCAGAGCCACAGTAAAGTGGAGTTTGTTC		348
Db	280	GGCGCGGCGCACGCGGCCACGGCGCTTAGGCACATGTGACCGGCAAGGTGGGGGTCTCATG		339
Qy	349	GGTACATCTGAGACCAAGGGGCGCATTAATGTGTGACCGGCATTTGCCAATGCCATTGGAC		408
Db	340	GCGACTTCTGTGCTCTGAGAGCCACCAATCTGGTGTACACCTGTGGCGCATGTGCAGATGAC		399
Qy	409	TTCGATGCCAATGATGATGATTAATCTGAGAGAGTGGGGCCGTGCAGATTTGGATGAGTCT		468
Db	400	TCAGTCCGATGTGCGCATTAATCTAGGCGAGGTCCGGCCAGTCTGATTCGGAGCGAATGCC		459
Qy	469	TTCCAGAGAAATGACATTTTGTGCGATCACTTACCGATCGTTAAGACCTCCATATGTGTA		528
Db	460	TTCCAGAGAGCGCGACATCTCCGCGATCAACAATGCGAATTAACAAGCATTAATTTCTGTGTT		519
Qy	529	CGTATGTCGGCGGATATGGCTCGATTTGTAATGAGCTTCCATCTGTGTAGACCGGT		588
Db	520	CGTGCCTGGCGACGATATTTCCCGGGTACTGGCTTAGGCTTTCACATCGGTGTCCGGA		579
Qy	589	CGTCCCGGCGCGTTTTGGATGATTAATCCCAAGGATGTGGCTTGAAGAATGTAGTAC		648
Db	580	CGCGCGGATGTGGTGTCTGTGTGATATCCCGAAGGATGTGTGACAGGCGCATGTCAATTC		639
Qy	649	ATTCCCTTCGACCCCGGTGACGTTAATCTACCGGGTTATGCCCCACGCGTTAAAGTAAT		708
Db	640	AG-----CTGGCACACGGAATAATGACCTTGGCCCGGTACAAAGCCGAACACCAACCGCAC		693
Qy	709	CCCCGACAAATTAATATGCGGGCATTTGCCAATGTTT3AGACAGGCGACAATACTCCTTCTTAC		768
Db	694	AACCGGCAGATCCGTGCGGACAGCCAACTGATGCTGACCGCGCAAGCCTGTGTGTAT		753
Qy	769	GTAAGGGGAGGGGCGATGCGCCGCAATGCCATGCCAGGTGCAGGAATTTGCCGAAGAAG		828
Db	754	GTCGCGCGGGCGCTTAATCCGCGGAGGAAGCGACCCAGAGGCTACCGGATTTGGCGAACTG		813
Qy	829	TTCCAGTTGCGCGGTAAACAACACCTGTATGAGGAATTTGGGGTTTTTGAAGAAAACATATCC		888
Db	814	AACCGGATCCCGGTGTGTTAGCAGCTGATGTGACATGTGCGCTTCCCCGAAAGCATATAC		873
Qy	889	CTTTCCGTGGTATGTTGGGTATSCATGTGACCGACCGATGTGCCAATTTGCCGTACGGA		948
Db	874	CAGAACTTGGGCATATGCCCGCATGCAAGGCACCTTGGCGGCGTGGGACAGCCTCTCAAGCC		933
Qy	949	TGTATTTGTTGATTTGACGTGGGGGCCCGTTTGCAGACCGGGGTAATCTGGCAAACTAGAC		1008
Db	934	AGCGATTTACTGATGTGCACTGGGCACTGTGCTTGTAGCGACGGGTGATCTGGCAAGCTGCAC		993
Qy	1009	GAATTTGTACCCGCGCCMAAGTAAATTCACATTGACATCGACCCGCGGAGGTGGGAAA		1068
Db	994	TCTCTTCGCGCCGACGCGCAAGTAAATATCATGCCACATCGACCCGCGGAGATCGGCAAG		1053
Qy	1069	AACAGGGTCCCGCATGTGTCCATTTGTGGGGATGTAGCCATGTTTAAACACTTTTG		1128
Db	1054	AACCGGACGCTGACGTGCCATTTGTGGGCGACGTCAAACTGTATTTGTAAGCTCATC		1113
Qy	1129	CAGCGGGCGCGGAAATGATTACCCCAACCCATCCCATACCAACCAGGCAATGTAAAT		1188
Db	1114	GCGATGTCCGCCCATTAAGAGGTCCCGGGCAACATGAAATAGACGACATGTGTGTATAT		1173
Qy	1189	CGCATTTGATTTGGCGGACCGATTAACCCCTCCAGGTGCCCACTATAGAGATATATTT		1248
Db	1174	TTGACGCTGTTTCGAAACGTAATCCGTTGAGCTACAGTCCACAAAGCGAGGCACTCTG		1233
Qy	1249	GCCCCCAGAGGTAGTACAGAAATTTGT--GCGCAGGCCCCGATGCTTACTACAC		1305
Db	1234	AGCCCGAGTACGTATTCGAAAAGCTGTGTAAATTTGTGGCTCCGAGGCGCTATATGTC		1299

QY 1306 ACCGATGTTGGGACCAACCAATGTTGGGCGCCAGT---TTTGAACATGGCCCCCGC 1362
 Db 1294 GGGGTTGATGTCACATCAGATGTGGGGCGGAGTTTCATCTTACGAGAGAGCGGT 1353
 QY 1363 CGATGATTTCCAGTGTGGCTTGGGATGATGGGCTTTGTTACTGCGCCATGGGA 1422
 Db 1354 ACTTGCTCACTCACTGCGGCTTGGGACCAATGGGATTTGCTATTCCGCGCCATGGGC 1413
 QY 1423 GCCAA-AGTGGAGTGGGGGACGAGCGGTCATTTCATCAGTGGAGATGCCAGCTCCAA 1481
 Db 1414 GCCAAGATTGCTCCCGGAGGCGCAAGTCTGGGGCATGACGCGGATGGCTGCTCCAG 1473
 QY 1482 ATGAATCTTCAGGAACCTGAGAACCTTACGATCAGATCAGTTAACTATTAT 1541
 Db 1474 ATGACCAACAGAGTGGCTACGTGCGCTATCGAGGGGCGCGCATCAAGTGGCGT 1533
 QY 1542 CTCATTAACGTTGGGAGGATGGTGGCTCAGTGGCAACAACCTTTCTACGAGAGCT 1601
 Db 1534 ATCAACACGCAACCTGGGATGGTGGGCAATGGCAAGCTTGTATTATCAGAGGCT 1593
 QY 1602 TATTCGCTTCTAATGATG-----TCCAGGGCATGGCAGACATTAATCTCTGT 1652
 Db 1594 TACTGCAAGACCGACCTGGCCACCACTCAGACCGCATTCGGACTTGTGAAGTGGCC 1653
 QY 1653 GAAGCTATGGCATCAAGGATTAATG---GTGGCAAGCGGGAAGATTGGCCCCGCG 1709
 Db 1654 GAGGGGTGGGTTGGCTGGATGGATGGCTGAGTGGAGAGAGAGTGTGATGTATC 1713
 QY 1710 ATCCGCAAAATGCTAGCCACATGATGATGATGATGATGATGATGATGATGAT 1769
 Db 1714 AATCAGGCGGAGGCGCATCAACAACCGCTGTGGTGAATGACTTATGTGGCGCGAT 1773
 QY 1770 GAAAACTGTACCTTATGATGATGATGATGATGATGATGATGATGATGAT 1823
 Db 1774 GCGCAGTGTGGCGCATGATGATGATGATGATGATGATGATGATGATGAT 1827

RESULT 6

US-10-282-122A-25560
 ; Sequence 25560, Application US/10282122A
 ; Publication No. US20040029129A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, Liangsu
 ; APPLICANT: Zamudio, Carlos
 ; APPLICANT: Malone, Cheryl
 ; APPLICANT: Haseibeck, Robert
 ; APPLICANT: Halseen, Karl
 ; APPLICANT: Zyskind, Judith
 ; APPLICANT: Wall, Daniel
 ; APPLICANT: Trawick, John
 ; APPLICANT: Carr, Grant
 ; APPLICANT: Yamamoto, Robert
 ; APPLICANT: Forsyth, R.
 ; APPLICANT: Xu, H.
 ; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
 ; FILE REFERENCE: ELITRA.034A
 ; CURRENT APPLICATION NUMBER: US/10/282,122A
 ; CURRENT FILING DATE: 2003-02-20
 ; PRIOR APPLICATION NUMBER: 60/191,078
 ; PRIOR FILING DATE: 2000-03-21
 ; PRIOR APPLICATION NUMBER: 60/206,848
 ; PRIOR FILING DATE: 2000-05-23
 ; PRIOR APPLICATION NUMBER: 60/207,727
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: 60/230,335
 ; PRIOR FILING DATE: 2000-09-06
 ; PRIOR APPLICATION NUMBER: 60/230,347
 ; PRIOR FILING DATE: 2000-09-09
 ; PRIOR APPLICATION NUMBER: 60/242,578
 ; PRIOR FILING DATE: 2000-10-23
 ; PRIOR APPLICATION NUMBER: 60/253,625
 ; PRIOR FILING DATE: 2000-11-27

PRIOR APPLICATION NUMBER: 60/257,931
 ; PRIOR FILING DATE: 2000-12-22
 ; PRIOR APPLICATION NUMBER: 60/267,636
 ; PRIOR FILING DATE: 2001-02-09
 ; PRIOR APPLICATION NUMBER: 60/269,308
 ; PRIOR FILING DATE: 2001-02-16
 ; Remaining Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 78614
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 25560
 ; LENGTH: 1866
 ; TYPE: DNA
 ; ORGANISM: Mycobacterium avium
 US-10-282-122A-25560

Query Match 19.4%; Score 379.8; DB 13; Length 1866;
 Best Local Similarity 54.3%; Pred. No. 3.5e-116;
 Matches 938; Conservative 0; Mismatches 757; Indels 34; Gaps 7;

QY 109 CTGAAAATGTTCTAGCCCAACGCAACCGGGCTTATATCTGATGATAGCTGAAA 168
 Db 100 CCGAAAGCATGGGCGCCGAAAGTTACTGTGCAACAATCGGTGATCCGATCGTGGAG 159
 QY 169 CGCCATGGGTCAACACATTTTGGCTATCCGGGGGGCAATTTTGGCCATCTATGAT 228
 Db 160 GAGCTGGCGTCAAGGTGATCTTCGCAATCCGGGCGCGGTGCTGCGGTCTATGAC 219
 QY 229 GAACTGATCCGCTTTGAAGCGGGGAAATTGACATATTTGGTGGCCATGAACA 288
 Db 220 CCGCTGTGACT-----CCAGAGCTCGGCACTGCTGATCCGCAAGAG 270
 QY 289 GAGCTCTCCATCGCGCGATGGGTATGCCAGCCACAGTAAAGTGGAGTTGTTTC 348
 Db 271 GGTGCGGTGACCGCCGCAAGCGCTACGGCAGCCACCGGCAAGGTGGTGTCTCAG 330
 QY 349 GGTACATCTGAGACAGGGGCACTAATGTTGATGACCGGATGCGCAATGCCATTGGAG 408
 Db 331 GCACTCTCGGTCCGGCGCCACCAACTGTGTACCGCGCTGGCCAGCCGATGAG 350
 QY 409 TCGGTGCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 468
 Db 391 TCGATCCCGTCTGCGCGCTGACCGGCAAGTGGAGACGACGCTGATCGGACCGAGCC 450
 QY 469 TTCCAGAAATTTGACATTTTGGCATCACCTTACGATGTTAAGACATCTATGATG 528
 Db 451 TTCCAGAGGCGCAACATCTCCGATCACCATGCGCATCAACAGCAAACTCTGATG 510
 QY 529 CGTAGTGGCGGATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 588
 Db 511 CGCTCGGCAAGAAATCCCGCAGGTGCTTGGCAGGCGTTCCACATCCGCGCTCGG 570
 QY 589 CGTCCCGGCGGTTTGAATGATATTTCCAGAGATGTGGCTTAAAGATGTGATG 648
 Db 571 CGCCCGGCGGATGCTGCGCATCCCAAGAGATGCTGCGAGGCGCAGTGCACATG 630
 QY 649 ATTCCCTGACCCCGGATGATGATGATGATGATGATGATGATGATGATGATGAT 708
 Db 631 AG-----CTGGCGCGCGGCACTCACCTGCGCGGTCAAGCGCACCAACCGCAC 684
 QY 709 CCCGACAAATTAATCGGATGATGATGATGATGATGATGATGATGATGATGATGAT 768
 Db 685 AGCCGGCAGATCCGCAAGCGCCAGCTGATGCGCGCGCGGCAACCGGTGCTGAT 744
 QY 769 GTAGGGGAGGGGATGCGCGGCAATGCCAATGCCAGTGAAGAAATTTGGGAAAG 828
 Db 745 GTAGGGGCGGTATCTATCCCGGGAGGCGCAGGACAGCTGGCGAATCTGCGAGCTG 804
 QY 829 TTCCAGTTCGCGTAAACAACACCTGATGGGATTTGGGCTTTTACGAAAACCATGCC 888
 Db 805 ACCGGATCCCGGTGATGACACAGCTGATGGCGCGCGCGGTCCCGACAGCCAGCCG 864
 QY 889 CTTTCGTGATATGTTGGTATGATGATGATGATGATGATGATGATGATGATGATGAT 948

D	865	CAGACACTGGGCATATGCCCGGGATGCAACGACACCGTCCGGGGGTGGCGGCCCTGCAGGCG	924
Q	949	TGTGATTTTGTGATTTGCAATGAGTGGGAGCCCGTTTCGACACCGGGTAACTGGCAAACTGAC	1008
D	925	AGCAACCTGTGATTCGGCTGGGTACGGGTTCGATGACCGGGTGAACGGGCAAACTGGAC	984
Q	1009	GAATTTGTAGCCCGGCGCAAGTAAATTCATTTGACATCGAACCGGCGGAGAGTGGGAAAA	1068
D	985	ACCTTCGCCCCGGGAAGCCAAAGTATCATACGCCCGACATCGAACCGGCGGAGATCGGCAAG	1044
Q	1069	AACAGGGCTCCGATGTGCCATTTGTGGGGATGTACGCCATGTTTGAACAGCTTTTG	1128
D	1045	AACCGGATGCGACGTCGTCGATCGTGGCGGACGTCAAGGCCGTCATGCGCAACTGGTC	1104
Q	1129	CAGCGGGCCCGGAGATTGGATTTAACCCACCATGCCATACACCCGACGCAATGTTAAAT	1188
D	1105	GAGATGCTGCGCCACGACGGGCGCCCGGCAATCTCGACATCGCCGATCTGTGGGCTTAC	1164
Q	1189	CGCATTTATCATTTGGCGGACCGCATTAACCCCTCCAGTGGCCCATAGAGATCTATT	1248
D	1165	CTCGACGACGCTCCAGTCCACCTATTCGGCTGAGTACGGGCGGACAGGAGCGGACGCTTC	1224
Q	1249	GCCCCCAGAGAGTAAAGTACAGAAATTTGTGCCAGGCG---CCGATGCTTACTACAC	1305
D	1225	GGGCCCGAGTACGTATCGAAGAACTGGGCGAGATCGCGGGCGGACGCCCTGTACGTG	1284
Q	1306	ACCAGTGTGGGACAAACCAATATGTGGCGGGCCAGT---TTTGAACATATGGCCCCGC	1362
D	1285	GCGGGCGGTGGTCAGCATCAGATGTGGCGGGCCCAATTCATCTCTCAAGAAACCGGCG	1344
Q	1363	CGATGGAATTTCAAGTCTGCGCTTGGGTACATGGGCTTTGTGTTAATCTGCGCCATGGGA	1422
D	1345	ACCTGGCTCAATTCGGCGGCTGGGACCAATGGGGTTCGCAATTCGGCGCGCATGGGC	1404
Q	1423	GCCAAATGGG-AGTGGGGGACGAGCGGTCAATTTGCATCACTGAGAGATGCGAGCTTCAA	1481
D	1405	GCCAAAGATGGGCGGTCCGGAGCGGAGGTGTGGGCGATCGAACGGGAGCGTGTTCACG	1464
Q	1482	ATGATCTTCAGGAATGGGAACCTTAGCCCATGTACGACATCCAGGTTAAATCTATTAT	1541
D	1465	ATGACCAACCGAGGAGCTGGCCACCTCGGCGCTGCAGGGCATTCGATCAAGTGGCGCTG	1524
Q	1542	CTCATATACGTTGGACGGGGATGTGTGTAGTGGCAACAACTTTCTACAGAAAGCT	1601
D	1525	ATCAACAAACGCAACTGGGCATGTGTGCGCCAGTGGCAGACCTCTGTTCTACAAAGAGCGG	1584
Q	1602	TATTC-----TGCTTAAACATGTCCCGAGGCATGCGCAACATTAATCTCTCTGT	1652
D	1585	TATTCGCAAAACGATCTGGCCACCACTCGCACCGCATCCCGGACTTCGTCAACCTGCGC	1644
Q	1653	GAACCCATAGGCATCAAGGGTAT---TACTGTGGCAAGCGGGAAGATTTGGCCCCGGCG	1709
D	1645	GAGCGCTGGGCTGCTGGATTCGCTTGCGAGGCTGGAAGACGTCGTGACGTGATC	1704
Q	1710	ATCCCGCAAAATGCTAGCCCAATATGTCCTGTGTGATGTGATGTGTGTGTCMAAAAAGAT	1769
D	1705	AAGCTGCGCGGGGAGATCAACGACCGGCGGTTGTGATCGACTTCATGTCGGGCGGAC	1764
Q	1770	GAAAACTGTAACTTATGATTTGCCCGGACATAGATATGCTCCCAATGCT	1818
D	1765	GCGCAGGTGTGGCGATGTGTGGCGCGGACCGGCAACGACGAGATCTC	1813

RESULT 7

US-10-282-1122A-41305
 Sequence 41305, Application US/10282122A
 Publication No. US2004002122A1
 GENERAL INFORMATION:
 APPLICANT: Wang, Liangsu
 APPLICANT: Zamudio, Carlos
 APPLICANT: Malone, Cheryl
 APPLICANT: Haselbeck, Robert
 APPLICANT: Olsen, Kari

```

APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: EITPA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/151,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PAM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 41305
LENGTH: 1722
/
/
/ ORGANISM: Vibrio cholerae
US-10-282-122A-41305

```

Query Match	19.4%;	Score 379.6;	DB 13;	Length 1722;
Best Local Similarity	53.9%;	Pred.No.3.9e-116;		
Matches 917;	Conservative	0;	Mismatches 164;	Indels 19; Gaps 6
QY	128	AACGGCAAAACCGGCGCTTATATCTCGATGATAGCTACCTGAAACCGCATGGGGTCAACACA	187	
Db	5	AAATGCTATCTGGCCGACAGATGATCGTCGAATCTGTGATCATGAAGGTGTGACCAAA	64	
QY	188	TTTTTGGCTATCCCGGGGGGCAATTTGGCCATCTATATGAATCTACCGCTTTGAAG	247	
Db	65	TCTTCGGATATCCCGCGGTTCAGTACTCGACATCTACGATCGCTTTCA-----GAAA	118	
QY	248	CGCGGGGGAAATTTAGACATATTTTGGTGCSCCATGAACAAGAGCTTCCCATCGCGCG	307	
Db	119	AAACACCAATTTAAACACGTTCTGTGTCTGTCAAGAAAGCGGCAACATATGGCTG	178	
QY	308	ATGGGTATCCAGAGCCACAGGTAAAGTGGGAGTTGTTCGGTACATCTGACACCGGG	367	
Db	179	ATGGCTATCCGGCGCCCACTGGAAAAACGGCGCGTAGTGCCTGTCTGTCAGTCCGGGG	238	
QY	368	CGACTACTTGGTGAACCGGCAATGGCCAAATGGCACTCGTGGTCCCATGATGTGA	427	
Db	239	CCACCAATACGGTGACTGGGATTCGAACCGCTCTATATGACATCGATTTCCGATATTTGA	298	
QY	428	TTTACTGGAGATGGGCGGTGCATGATGGTGAACGATGCTTTCACAGAAATTGACATTT	487	
Db	299	TCTCCGGTAACTACCACTAATCTTATTTGGTAAACGATGATTTCAAGAGTGGATATCG	358	
QY	488	TTTGGATCACTTACCGATCGTATGAACCTCTATGGTGAACGATGGCGCGCATATGG	547	
Db	359	TGGCGTGTCTGTCCATCGTTAAACACAGCTTTTGGTGGAAAAAACCGGAGATATTC	418	
QY	548	CTCGATTTTACTAGGCGTTTCACATCTTGCTACACCGGATCGTCCCGGCGGTTTTGA	607	

Db	419	CTGAACCATCAAAAAGGTTTATATGCGCTCAACGGGGGCTCTGCTCTCTATAGTCA	478
QY	608	TGGATATTCACAGAGATGTGGGCTTAGAAGAAATGTAGTACATTCCTCCCTGACCCCGGTG	667
Db	479	TTGACGTGCGCAAAAGATGTGATGAGCCCGCTTAATAAGTGTGCTTATGAGTATCCAGAAA	538
QY	668	ACGTTAATCTACCCGGTATCGCCCCACGGTTAAAGGTATATCCCGACAAATTAATGGCG	727
Db	539	CCATTAAATATGGCTCTTACAACTCAACGACGGCTGGCCACAAAGACAGATCAAAAAAG	598
QY	728	CATTGCAATTGTTGAGCAGGCGCAAAATCCCTTGCTCTACGTAGGGGAGGGGCGATCG	787
Db	599	GGTTACGGCGCACTGCTCGATGCGAAAAAGCCAGTACTGTAGCTCGGTGGCGGGCGGATTA	658
QY	788	CCGCAATAGCCCATGCCAGGAGGAGGAAATTTGGCCGAAAGAGTTCCAGTTGCCGCTAACAA	847
Db	659	TTTGCTAAATGACAGATAGCAAAATTCGCCCACTTGACAGAGCTTTAAATCTGCCGTGTTA	718
QY	848	CCACCCGTATGGGAAATTTGGGCTTTTGAAGAAACATCCCTTCGGTGAGGTATGTTGG	907
Db	719	GTACTTTATGAGGATTGGGCGTTTTTTCGGGCACTCAATAAAACCTCACCTGGATGCTTG	778
QY	908	GTAATGCAATGGCCACCGCTATGCGCACTTTGCCGTAGCGAATGTATTTGTTGATTGCA	967
Db	779	GCATCCATAGGTGTGTATGAGGCCAATATGGCCATCACAATGCAATGATCGATTTTGGGG	838
QY	968	TGGGGGCCCGTTTCGACGACCGGGTAACTGGGCAACTGAGCAATTTGTGATGCGCGCA	1027
Db	839	TAGGTGTGCGCTTTATATGACGCTACGACTAACAACTTGAAGAAATGCTGCCCAACGCA	898
QY	1028	AAGTAATTCACTTGACATTCGACCCGGCGGAGTGGGAAAAACAAGGCTCCCGATGTGC	1087
Db	899	AAATCATCATATCATATTTGACCCGTCATGATTTCTTAAAAAGTGAAGTCGATCTGCG	958
QY	1088	CCATTGTGGGGAGTGTACGCCCATGTTTTAGAACAGCTTTTGACAGGGGCCCGGGAATTGG	1147
Db	959	CGATTGTGTTCTGCGCCACCAAGATTGGAGTGCGCATCTCAAGGTTTCTGAAAGGAGCG	1018
QY	1148	ATTACCCCAACCCATCCCATACCAACCCAGGCGATGTGTTAAATGCAATTGATCATATGGCGGA	1207
Db	1019	CTGAGCGCAATGATGCGGCGAGATTTGATAGTGTGTGAGATGAGATCCAAAGTGTGCTG	1078
QY	1208	CCGATTAACCCCTCCAGGTGCCCACTATGAGGATATCTATTGCCCCCAAGA--GGTAG	1267
Db	1079	ATCGCAATGTTTGGCTTATGAAACCTACAGCAACGAGTTAAGCTCAGCAAGATTTG	1138
QY	1265	TACACGAATTTGTTGCGCAGGCGCCCGAGTCCCTACTACACACGATGTGGGACAAACCC	1324
Db	1139	AAAGGCTATTAATAATCAAGGAAGTTAAAGCAATTTTGCCCTCGGATGTGGGCGACACC	1199
QY	1325	AAATGTGGGCGCGC--CCAGTTTGTGAACAATGCCCCCGCGCATGATTTTCCAGTGTG	1381
Db	1199	AAATGTTTGGCGCACTACTATACGTTTGATTAACCGCGTCAAGTGAATTAACCTGTGGCG	1255
QY	1382	GCTTGGGTAGCATGGGCTTTGTTTAACTGCGGCATATGGAGCCAAAGTGGAGTGGGGG	1441
Db	1259	GGCTAAGGCAACATGGGTTTGGTTTAACTGCGGAGATGGGCGTGAATTTGGCAAGCCAG	1318
QY	1442	ACG-AGCGGTATTGTCATGATGAGAGATGCGAGTTCMAATGATCTTCAGGAACTGG	1500
Db	1319	AAGAAGAGATGCTGCTGCTGACCGGAGATGGAGATTCAGATGAATTTCAAGAGCTAT	1378
QY	1501	GAACTCTAGCCCATGTAGCAATCCAGGTTAAACTATTAATTCATAAATACGTTGGCAGG	1560
Db	1379	CGACCGCACTGCATGACACATTCACGATTAATTAACCTCAACACCGTTTCTTTG	1438
QY	1561	GGATGTGCTGCTGAGTGGGCAACAACTTCTACGAAGAGGTATCTGCTTCAACATGT	1620
Db	1439	GAATGTGAACAGTGGCGAGATATATTATTTCAAGGTGCTCATCTTAATCATATCAATGA	1498
QY	1621	CCCAAGGCGATGCCAGACATTAATCTCTGTGTGAAGCTATGACATCAAGGTAATTAAGT	1680

Qy	Db	Sequence	Score	Length	Indels	Gaps	
Qy	Db	1499 GTTGTG---TGCGGAGATTTTGCGGCGATTCGCAGAAAGCTTATGCTCATGTTGGGATCCGTA	1555				
Qy	Db	1681 TCGCGAAGCGGGAAGATTGTGCCCCCGCGATCGCCGAAATGCTAGCCGACATGTCCTG	1740				
Qy	Db	1556 TTTCCTACCCCTGAGAGACTGCMAAGCTGGGGCTGGGAAAAAGCGTTAGCATGAAAGATCGTC	1615				
Qy	Db	1741 TGTGTG---ATGAGATGTGTGTGTCAAAAAGATGAAAACTGTATACCTATGATTTGCCCGC	1797				
Qy	Db	1616 TGTGTTTGTTCGATTAATTAATGTCATGAAACTGAGCAGCTTTACCAATGCAATCAAG	1675				
Qy	Db	1798 GCATAGTAAATGCCCAATG	1817				
Qy	Db	1676 GCGAAGGATGACCAAAATG	1695				
<p>RESULT 8</p> <p>US-10-282-122A-11812</p> <p>Sequence 11812, Application US/10282122A</p> <p>Publication No. US20040029129A1</p> <p>GENERAL INFORMATION:</p> <p>APPLICANT: Wang, Liangsu</p> <p>APPLICANT: Zamudio, Carlos</p> <p>APPLICANT: Malone, Cheryl</p> <p>APPLICANT: Haselebeck, Robert</p> <p>APPLICANT: Ohlsen, Kari</p> <p>APPLICANT: Zyskind, Judith</p> <p>APPLICANT: Wall, Daniel</p> <p>APPLICANT: Trawick, John</p> <p>APPLICANT: Carr, Grant</p> <p>APPLICANT: Yamamoto, Robert</p> <p>APPLICANT: Forsyth, R.</p> <p>APPLICANT: Xu, H.</p> <p>TITLE OF INVENTION: Identification of Essential Genes in Microorganisms</p> <p>FILE REFERENCE: ELITRA.034A</p> <p>CURRENT APPLICATION NUMBER: US/10/282,122A</p> <p>CURRENT FILING DATE: 2003-02-20</p> <p>PRIOR APPLICATION NUMBER: 60/191,078</p> <p>PRIOR FILING DATE: 2000-03-21</p> <p>PRIOR APPLICATION NUMBER: 60/206,848</p> <p>PRIOR FILING DATE: 2000-05-23</p> <p>PRIOR APPLICATION NUMBER: 60/207,727</p> <p>PRIOR FILING DATE: 2000-05-26</p> <p>PRIOR APPLICATION NUMBER: 60/230,335</p> <p>PRIOR FILING DATE: 2000-09-06</p> <p>PRIOR APPLICATION NUMBER: 60/230,347</p> <p>PRIOR FILING DATE: 2000-09-09</p> <p>PRIOR APPLICATION NUMBER: 60/242,578</p> <p>PRIOR FILING DATE: 2000-10-23</p> <p>PRIOR APPLICATION NUMBER: 60/253,625</p> <p>PRIOR FILING DATE: 2000-11-27</p> <p>PRIOR APPLICATION NUMBER: 60/257,931</p> <p>PRIOR FILING DATE: 2000-12-22</p> <p>PRIOR APPLICATION NUMBER: 60/267,636</p> <p>PRIOR FILING DATE: 2001-02-09</p> <p>PRIOR APPLICATION NUMBER: 60/269,308</p> <p>PRIOR FILING DATE: 2001-02-16</p> <p>Remaining Prior Application data removed - See File Wrapper or PALM.</p> <p>NUMBER OF SEQ ID NOS: 78614</p> <p>SOFTWARE: PatentIn version 3.1</p> <p>SEQ ID NO 11812</p> <p>LENGTH: 1875</p> <p>TYPE: DNA</p> <p>ORGANISM: Burkholderia cepacia</p> <p>US-10-282-122A-11812</p>							
Qy	Db	136 ACCGGGGCTTAATCTGATGATAGCCTGAAAGCCCATGGGCTCAAAACATTTTGGC	195	19.3%;	Score 378.8;	DB 13;	Length 1875;
Qy	Db	178 ATCGGGCGGCAACCGTGTCTCATGAAAGGACGTCGCGAGCAAAACGTCGATTCATCTGGGGC	237	Best Local Similarity 54.7%;	Pred. No. 7.7e-116;	Mismatches 939;	Conservative 0;
Qy	Db	136 ACCGGGGCTTAATCTGATGATAGCCTGAAAGCCCATGGGCTCAAAACATTTTGGC	195				
Qy	Db	178 ATCGGGCGGCAACCGTGTCTCATGAAAGGACGTCGCGAGCAAAACGTCGATTCATCTGGGGC	237				

QY 196 TATCCCGGCGGCGCAATTTTGGCCCATCTATGATGACCTGTACCGCTTTGAAAGCGGCGGG 255
 Db 238 TACCCCGGCGGCGCTCGTTCTCTACATCTACAGAGAGTGTACA-----AGCAGGAC 288
 QY 256 GAAATTGAGCATATTTTGTGCGCCATGAAACAGAGACTTCCATGCGGCGGATGGGTAT 315
 Db 289 AAGATTGACGACGTGCGTGGCCATGAAACAGCGGCGCTGTGACGAGCGGATGCGTAT 348
 QY 316 GCCAGAGCCAGAGTAAAGTGGAGTTTGTTCGGTACATCTGAGCCAGGCGGAGACTAAC 375
 Db 349 GCGGCTTCCACCGGCAACGTGCGGCTGTCTCTGTGACGTGCGGCGCGGCTGTGACCAAT 408
 QY 376 TTGGTACCGGCAATTTGCCAATGCGCATTTGGACTCGGTGCCCATGTGTGATTACTGGA 435
 Db 409 GCGGTGACCGGCAATGCAACCGGCTTACATGATTCGATCCGATGCTGTATCAGCGGC 468
 QY 436 GAGGTGGCGCGTGCATGATGTGTAGAGATCTTTCCAGAGAAATGACATTTTGGCATC 495
 Db 469 CAGGTGCGGACCGCGGATGCGTCAAGATGCGTTCCAGAGATGCGACACCGTGGCATC 528
 QY 496 ACCTTACCGATGTTAAGACATCTTATGTGTAGACGTGTGCGGCGGATGCGTCTGCAT 555
 Db 529 ACGGTCGCGTGTGAAGCAACCTTCTCGTGAAGACGTGCGCGACCTGCGGAAAC 588
 QY 556 GTTACTGAGGCTTCCATCTTGTAGACACCGGTGTGCGGCGCGGCTTTGATGCAAT 615
 Db 589 GTCAAGAGCGCTTATATGCGCGGACCGGCGGCGGCGGCTGTGTATGACATC 648
 QY 616 CCGAAGAGATGTGGGCTTAAAGAAATGTGATGACATTCCTCGACCCCGGTGACGTTAT 675
 Db 649 CCGAAGAGCATCTGGAAGACCGGTGCGAAGTACAGACGATCA-----GAGGTGTG 702
 QY 676 CTACCGGCTTATGCGCGGCGGCTTAAAGTATTCCTCGACCAATTAATGCGGCTTGA 735
 Db 703 CTGCGTTGTCAACCCCGGTGACAGAGGCGGCTTGTGCGTCAATCCGCAAGCGGTGTG 752
 QY 736 TTGTTGAGAGCGGCGCAAAATCCCTGTCTACGTAGGCGGAGGCGGATGCGGCAAT 795
 Db 763 CTGCTGTGAGCGGCGAAGCGTCCGTACATCAACGCGGCGGCGGCAATCTCTGCGGAC 822
 QY 796 GCCATGCGGCGAGTGCAGAAATTTGCGGAAAGTTTCCAGTTGCGGCTTGAACACACCTG 855
 Db 823 GCGTGCAGTGAATGAAACAGTTGCGGACCTGTCTGCGTCAACCGGCTGCAAGACCTG 882
 QY 856 ATGGGAATTTGGGGCTTTTGAAGAAACATCCCTTTCGAGGGTATGTTGGGATCAT 915
 Db 883 ATGGGCTTGGCGGCTATGCGGCTGCGGCAAGAAATTCCTCGGCAATGCTGCGGATCAC 942
 QY 916 GCGCAACCGCTATGCGCAATTTGCGGTGACGCAATGTATTTGTTGATGAGTGGGCGC 975
 Db 943 GCGCACTTACGAAGCGAATGCGGATGCGACATGCGACGTCGTGATCGGATCGGCGCC 1002
 QY 976 GATTTCGAGACCGGCTTACGCAACTAGCAATTTGCTACCGCGC---CAAGTA 1032
 Db 1003 GCTTCGAGACCGGCTGATCGGCGACCGGCGGCACTTCGCTGCGCGCGCGGCAATC 1062
 QY 1033 ATTGACATTTGACATGACCCGCGGAGGTGGGAAAAACAAGGCTCCCGATGTGCCAT 1092
 Db 1063 ATTGACATGACATGACCGGCTGCGTGCATCTGAAAGCGGTGAAGTTCGACATCCGATC 1122
 QY 1093 GTGGGAGATGTACCGCATGTTTGAACAGCTTTTGAAGCGGCGCGCGGAAATGATTA 1152
 Db 1123 GTGCGCGAGTGAAGAGTGTCTGAAGAGCTGATGACGAGTTGCGAGCGCGGACAT 1182
 QY 1153 CCGACCCATCCCATATACACCGGAGCATGTTAATGCAATGTATGTTGGCGGACGAT 1212
 Db 1183 GCGCGCGGACACCGAGCGCTGCGGCAATGTGGAAGACATCGAAGGCTGGCGGCGAAG 1242
 QY 1213 TACCCCTTCAGGTGCGGCGGCTATGAGGATCTATTGCGCGCGCGGAGGTGATACGAA 1272
 Db 1243 GACTGCGTGAAGTACGACCGGAAAGCGAGATCATCAAGCGCGAGTGTGTGAGAG 1302
 QY 1273 ATTGGTGGCGGCGCGCG---ATGCTACTACACCAACGATGTGGAGCAACCAATG 1329

Db 1303 GCGTGGAGCTGACCGAGCGGCAATGCGTGTGCTGTGAGCGTTCGCGGACACCAAGATG 1362
 QY 1330 TGGGCGGCGCAGTTT---TGAAATGAGCGCGCGCGATGATTTTCCAGTGTGCTTG 1386
 Db 1363 TGGGCGGCGCAGTTTACCGTTTGAACAGCGCGGTGTGATCACTCCGCGGCGCTC 1422
 QY 1387 GGTACGATGCGGCTTTGTTTACCTTCCGCGCATGAGAGCGCAAGTGTGGAGTGA 1445
 Db 1423 GCGACGATGCGGCTTGGCGCTGCGCGCGGATGAGCGGTCAAGATGCGGCGGATC 1482
 QY 1446 GCGGTCAATTTGATGAGTGAAGTATGCGGCTTCCAAATGATTTTCAAGAACTGGAAAC 1505
 Db 1483 GAGGTGTGTGATACGAGGAGAGGCTGATTCAGAGTGTATTCAGAACTGTGAC 1542
 QY 1506 CTAGCCAGTACGACATCCAGGTTTAACTTATTTCTCAATTAACGTTGCGAGGAGT 1565
 Db 1543 TGCCTGAGTACGACCGCGGTGAAGATCAATTCGCTGAACACCGCTACCTCGGATG 1602
 QY 1566 GTGCTGTGTGCGCAACAACTTTTCAAGAAAGCTTATGCTTTCAATGATGCTCAG 1625
 Db 1603 GTCCGCGAGTGCAGACAGATGAAATACGAGCGCTATTTGCAATTCGATCAATG---GAT 1659
 QY 1626 GGCATGCGGACATTAATCTCTCTGTGAGCTTATGCGATCAAGGATTAATCTGTGCGC 1685
 Db 1660 GCGCTGCGGATTTGTGTAAGCTGCGGAGCGTACGCGCATGTCGCGATCGGATCGAA 1719
 QY 1686 AAGCGGAAAGATTGGCGCGCGCGATGCGCGAATGCTAGCCGCAATG---GTCTGTG 1742
 Db 1720 AAGCTTGTGATGTGAGCGCGCGCGCTGAAGAGCGCGCTGAAAGACCGACCGTG 1779
 QY 1743 GTGATGATGTGTGTGCAAAAAGATGAAATCTTATGCTTATGCTTATGCTTATG 1802
 Db 1780 TTTCGTGACTTCCAGACCGATCCGACCGAAACGTCGTGCGGATGTGACAGCGGCGAAG 1839
 QY 1803 AGTAATGCCCAATCTAGGTTTACCGGAGTGGCG 1838
 Db 1840 GGCATCAGGAGATGCTGCTGATCGGAAATGTC 1875

RESULT 9
 US-09-815-242-7947
 ; Sequence 7947, Application US/09815242
 ; Patent No. US20020061569A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Haselbeck, Robert
 ; APPLICANT: Ohlsen, Kari L.
 ; APPLICANT: Wall, Daniel
 ; APPLICANT: Zyskind, Judith W.
 ; APPLICANT: Trawick, John D.
 ; APPLICANT: Carr, Grant J.
 ; APPLICANT: Yamamoto, Robert T.
 ; APPLICANT: Xu, H. Howard
 ; TITLE OF INVENTION: Identification of Essential Genes in
 ; FILE OF INVENTION: Prokaryotes
 ; FILE REFERENCE: ELITRA.011A
 ; CURRENT APPLICATION NUMBER: US/09/815.242
 ; PRIOR FILING DATE: 2001-03-21
 ; PRIOR APPLICATION NUMBER: 60/191,078
 ; PRIOR FILING DATE: 2000-03-21
 ; PRIOR APPLICATION NUMBER: 60/206,848
 ; PRIOR FILING DATE: 2000-05-23
 ; PRIOR APPLICATION NUMBER: 60/207,727
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: 60/242,578
 ; PRIOR FILING DATE: 2000-10-23
 ; PRIOR APPLICATION NUMBER: 60/253,625
 ; PRIOR FILING DATE: 2000-11-27
 ; PRIOR APPLICATION NUMBER: 60/257,931
 ; PRIOR FILING DATE: 2000-12-22
 ; PRIOR APPLICATION NUMBER: 60/269,308
 ; PRIOR FILING DATE: 2001-02-16
 ; NUMBER OF SEQ ID NOS: 14110

```

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7947
; LENGTH: 1725
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1725)
; US-09-815-242-7947

```

```

Query Match      19.1%; Score 373.8; DB 9; Length 1725;
Best Local Similarity 54.6%; Pred. No. 3.5e-114;
Matches 886; Conservative 0; Mismatches 712; Indels 25; Gaps 6;

```

```

QY 135 AACCGGCGCTTATCTGATGATAGCTGAAAGCCATGGGGTCMAACACATTTTGG 194
Db 12 ATCTGGCGCTGAATGCTGCTCCGCTCGTTGGCGACAGAGCGTTAAGTACATCTACGG 71
QY 195 CTATCCCGCGCGGCAATTTTGGCCATCTATGATGAACCTGTACCGCTTTGAAGCGCGGG 254
Db 72 GTACCCCGGCTGGTCCCTCTGCTATCTACGACCGCTCTTC-----AAGGAGCA 122
QY 255 GGAATTTAGCATATTTTGGTGGCGCATGAACAAGACCTTCCCATGGCGGATGGGTA 314
Db 123 CGACGTGACCCATCTGCTGGTGGCGCATGAAGAGCGGCTACCCATATGGCCGAGCTA 182
QY 315 TGCCAGAGCCACAGGTAAAGTGGAGTTTGTTCGCTACATCTGAGCCGAGGCGACTAA 374
Db 183 CGCCCGCGCCACCGGCAAGCCCGGTGCTGCTGTGTACTTCGCGCCGCGGCGAGCCAA 242
QY 375 CTGTGTGACCGGCAATTTGCAATGACCTGCTGGTCCCATGCTGTGATTTACTGG 434
Db 243 CGCCATCACCGGTATCGCCACGCGATACATGACTCCCATCCGATGATCTGTGCGG 302
QY 435 AGAGTGGGCGCTGCGCATGATGTGAGCGATCTTCCAGAAATTTGATTTTGGCAT 494
Db 303 CGAGGTGGCGAGCAATGCTGTGATCGATGCTTCCAGAAACCGCATGCTGGAT 362
QY 495 CACCTTACCGATCTTTAAGCACTCTTATGTGTGATGATGAGGCGGATATGCTCGAT 554
Db 363 CTCCCGCGCGATCTGGAAGACACGCTTATCATCAAGCATCCCTCGGAAATTCGCGAGT 422
QY 555 TGTTCGTGAGCTTCCATCTTGTAGACACCGGCTGCTCCCGGCGGCTTTGATGAT 614
Db 423 GATCAGAAGGCTCTTCTACTGCGGCACTCGGCTGCTCGGCGGCGGCTGTGAGAT 482
QY 615 TCCCAAGATGTGGCTTGAAGAAATGAGATCAATTCCTCGACCCCGGTGACGTAA 674
Db 483 TCCGAAGGACATGAGGCGACCGGACGCAAAATTCGATATTCCTTCGAGAAAGGTCA 542
QY 675 TCTACCGGCTTATGCGCCCGACGCTTAAAGTATCCCGACAAATTAATGGGCAATGCA 734
Db 543 GTTGCGTTGTCACACCGCGCGCTTGCCTGCTCACTCGGAGACATCCGAAAGCGCGCA 602
QY 735 ATTGTGAGAGCAGCAGAAATCCCTGCTACGTAGGAGGAGGCGGATCGCGCCAA 794
Db 603 GATGCTCTGCGCGCGCAAGCGCGCGGTGTGTATTCGCGCGCGCGCGGATCATGCGGTA 662
QY 795 TGCCCATGCCAGGTGCGAGAAATTTGGCGAAAGGTTTCCAGTTGCCGCTGAACAACCTCT 854
Db 663 TGCCCGCGCGCGCTGACCGAGCTGGCCGATGCTCAACCTGCGCGGTGACCAACACCTT 722
QY 855 GATGGGAATTTGGGCTTTTGAAGAAACCATCCCTTCCGATGGGATGTTGGGATGCA 914
Db 723 GATGGGCTCTCGCGGCTATCCCGGCGAGACCGGATGTTCTCGGATGCTCGGAGATCA 782
QY 915 TGCCCAACGCTATGCGCAACTTTCGCGTACGCGAATGTGATTTGATGAGTGGGCGC 974
Db 783 TGCCAGCTTCAACCGCAACTGCGGATGATCAACGACGATGATCTCGCGGTGCGGCGC 842
QY 975 CGGTTTCAGCAGCGGCTGATGCGCAACTGAGAGAAATTTGCTAGCCGCGCAAGAT 1034
Db 843 GCGTTTCAACCGGCTGATCAACGATGCGGCAAGTTTGTGCCCGAAAGCGCAAGATCAT 902

```

```

QY 1035 TCACTTACATGACCCCGCGAGGTTGGAAAAAACAAGGCTCCGATGTGCCATTTGT 1094
Db 903 CCACATGACATGATCCGCGCTGCTGATTTCCAAAGACATCAAGGCGGACATCCGATCGT 962
QY 1095 GGGGATGTAGCGCATGTTTAAAGACAGCTTTTGCAGCGGCGCGGGAATTTGATTAAC 1154
Db 963 CGGCGCGGTGGAAGAGGTTCTCAACGAATGCTGGATGCTGAGGAAATCCGCGAGAC 1022
QY 1155 CACCATCCCAATCAACACCGAGCATGTGTTAAATCGCATGATCAATGSC--GACCGA 1211
Db 1023 CCGGAACAGAGTCCCAAGGCTCTGCTGTGAAGCAGATGACGAGTGGCTGTGAACCG 1082
QY 1212 TTACCCCTCCAGGTGCCCATATGAGGTAATATTTGCCCCCGAGAGGTAGTACACGA 1271
Db 1083 TGGTGTCTCCGTAGACAAAGGCGACGCGACATCATCAAGCCGACGCTCATCGA 1142
QY 1272 AATTGTGCGCAGGCCCC-----CGATGCTTACTACACACCGATGTGGACACACGA 1325
Db 1143 GACGCTTACAGATTAACCATGCGATGCTTATCACTCCGATGTGCGCAGACCA 1202
QY 1326 AATGTGGGCGCGCCAGTTTGA---CAATGCCCCCGCATGTGATTTCCAGTCTGG 1382
Db 1203 GATGTTGCTCTGCACTACTACAAGTTCAACAAAGCCCAATGCTGTGATCAATCCGCGG 1262
QY 1383 CTGGGTACGATGGGCTTTGTTTACCTGCGCCATGAGGAGCCAAAGTGGAGTGGGGA 1442
Db 1263 TCTGGACCATAGGGCTTGGCTTCCCGGCGGCGATGGGATCAACCTCACTCCCGGA 1322
QY 1443 CGA-GCGGTATTTGCATCAGTGAAGATGCCAGCTTCCAAATGATCTTCAGGAATGGG 1501
Db 1323 CGAGATGCTGCTGTGTCGCGGCGAGGCGAGTATCCAGTAAATCCAGAACTGTC 1382
QY 1502 AACCTTAGCCAGTACGACATCCAGGTTAAACTTATTTCAATAACGTTGGCAGGG 1561
Db 1383 TACCTGCTGCAATGACACTGCGGTGAAGATGCTCAACTGAACAATGATCTCCCTGGG 1442
QY 1562 GATGTGCTGCTAGTGCACAACTTTTCAAGAAAGCTTATTTGCTTCTTCAACATGTC 1621
Db 1443 CATGTTCGCGCAGTGCAGAGCAATGCAATCAACACCGTTATTCGATTTCCATCAATG-- 1500
QY 1622 CAGGGGATGCGCAATTAATCTCCTGTGTAAGCTATGAGATCAAGGTATTAATCTG 1681
Db 1501 -GATGCTCTGCGGACCTTGTCAAGCTTGCAGAGCTTACGGGCAATGTGCGCAT 1559
QY 1682 GCGCAAGCGGGAATTTGGCCCCCGGCGATCGCCGAATGCTAGCCCAATGATGCTGT 1741
Db 1560 CACCGACCTGAAGATCTGAAGCGGAATGAGGAGGCGTTGCGCATGAAGATGCGCT 1619
QY 1742 GGT 1744
Db 1620 GGT 1622

```

```

RESULT 10
US-10-282-122A-30524
; Sequence 30524, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Malone, Cheryl
; APPLICANT: Zamudio, Carlos
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TIME OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A

```

```

CURRENT APPLICATION NUMBER: US/10/282,122A
;
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PAM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO: 30524
; LENGTH: 1725
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-10-282-122A-30524

Query Match      19.1%; Score 373.8; DB 13; Length 1725;
                  54.6%; Pred. No. 3.5e-114;
Matches 886; Conservative 0; Mismatches 712; Indels 25; Gaps 6;

QY 135 AACGGGGCTTATTCCTGATGATAGCTGAAAGCGCATGAGGGTCAACACATTTTGG 194
DB 12 ATCTGGCCCTGAATGTCCTCCGCTGTCGCGAAGAAAGCCTTAAATCACTACGG 71
QY 195 CTATCCGCGCGGCGCAATTTTGCCTATATGATGAACCTGACCGCTTTGAAGCGCGG 254
DB 72 GTACCCGGGTGTGCTCTGCAATCTAGACGCGCTTTC-----AAGAGCA 122
QY 255 GGAATTTAGCATTTTGTGTGCGGCATGAAACAAGAGCTTCCCATCGCGCATGTGTA 314
DB 123 CGAGGTGACCAACATCTCGTGTGCGCAAGAACAGCGGCTAACATATGCGCGACTA 182
QY 315 TGCCAGAGCCACAGGTAAAGTGGAGTTTGTTCGTAACATCTGACACAGGCGACTAA 374
DB 183 CGCCCGCGCCACCGGCAAGCCCGGTGTGTGTGATCTTCGCGCCGCGGCGGACCA 242
QY 375 CTGTGTGACCGGCATTTGCCATTTGCCATTTGGACTCGGTGCCCATGTGTGTGAT 434
DB 243 CGCATACCGGTATGCGCACCGCATACATGAGACTCCATCCGATGTGATCTGTGG 302
QY 435 AGAGGTGGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 494
DB 303 CCAAGTGGCGCAACATGTCGTGTCGATGTCGATGTCGATGTCGATGTCGATGTC 362
QY 495 CACTTACCGATCGTTAAGCACTCTATGTGTAGTGTGTGTGTGTGTGTGTGTGT 554
DB 353 CTCCGCGCGCATCGTGAAGCAAGCTTATCATCAAGATCCCTCGGAAATTTCCGAG 422
QY 555 TGTATCTGAGCTTTCATCTTGTGTAGACCGGTGTGTGTGTGTGTGTGTGTGTGT 614
DB 423 GATCAAGAGGCTTTCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 482
QY 615 TCCCAAGATGTGGCTTTAAGAAATGTAGTATATTCCTGACCCCGGTGTGACTTAA 674
DB 483 TCCGAAGACATGTGGCGACCGACGCAAGAAATTCGAATATTCATCCGAAGAGTCAA 542
QY 675 TCTACCGGTTATGCGCCCAAGTTAAAGTAAATCCCGCAAAATTAATGCGGATTCGA 734

```

RESULT 11
US-09-815-242-5925
Sequence 5925, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Cair, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO: 5925
LENGTH: 1815
TYPE: DNA
ORGANISM: Escherichia coli
FEATURE:
NAME/KEY: CDS
LOCATION: (1)...(1815)
US-09-815-242-5925

Query Match 18.8%; Score 368.6; DB 9; Length 1815;
Best Local Similarity 55.3%; Pred. No. 2,1e-112;
Matches 805; Conservative 0; Mismatches 634; Indels 16; Gaps 4;

QY 172 CATGGGCTCAACACATTTTGGCTATCCCGCGGCGCAATTTGGCCATCTATGATGA 231
DB 139 CAGGGCGTTAAACAAGTATCGGTTATCCCGAGCGCGACGCTTGATATTTATGATGCA 198
QY 232 CTGTACCGCTTGAAGCGGCGGGAATGAGCATATTTTGTGCGCCATGACAGAGA 291
DB 199 -----TTGCATACCGGTGGGTGATTTGATCATGATATGTTTGTGATGAGCAGCG 249
QY 292 GCTTCCCATGCGCGGATGGTATGCCAGAGCCACAGTAAAGTGGAGATTGTTCCGT 351
DB 250 GCGGTGATATGCGCGATGGCTGCGCGCGCGACCGGGGAAGTCGGGTGCTGTGTA 309
QY 352 ACATGTGACCGAGGCGGAGCTAATTGTGACCGGCGATTTGCCANNGCCATTTGGAATCG 411
DB 310 ACGTGGGTCCAGGGGCGACCAATGCCATTTACTGCAATGCCACCGCTTATATGATCC 369
QY 412 GTGCCCATGTGTGATTAATGAGAGGTGGCGGTGCGCATGATGTGAGCGATCTTTC 471
DB 370 ATTCCATTAGTTGTCTTCCGGGCGAGGTAGCAGACCTGTTGATAGTTAGATGCTTT 429
QY 472 CAGGAATATGACATTTTGGGATGACCTTACCGATCGTTAGCACTCTCTATGTGTACGT 531
DB 430 CAGGAGTCCGACATGTGTGGGATTTTCGACCGGCTGTTAAACAAGTTTCTGTTAAG 489

QY 532 AGTCCGCGGATATGCTCGCATTTGTACTGAGGCTTTTCATTTGCTAGACACCGGTGCT 591
DB 490 CAACCGAACAATTCGCGAGGTGCTGAAAAAGGCTTTCTGCTGGCGGCAATGGTCCG 549
QY 592 CCCGGCGGCTTTTGAATGATATTTCCCAAGAGATGGGCTTTAAGAAATGTGATCAT 651
DB 550 CCAGAACCAATGATCGTTGATTTTACCGAAGATATTTTATATCCGGGAACAAATTA 609
QY 652 CCCCTGACCCCGGTGACGTTAATCTACCGGTTATGCCCAAGTTAAAGTTATCC 711
DB 610 TATGCTGGCGGAGTCGTCAGTATGCCGTTTCAATATCCCACTACTACCGGACAT 659
QY 712 CGACAAATTAATCGGATTTGCAATTTGTGAGACAGCCAGAAATCCCTTGCTCTAC 771
DB 670 GGGCAATTAAGGTGCTCTGCAACGCTGTAGCGGCAAAAAACCGGTTGTTACGTA 729
QY 772 GGGGAGGGGCGATCGCCCGCAATGCCATGCCAGGTGACAGAAATTTGCGAAGATTC 831
DB 730 GCGGTGGGCAATTCAGCGGCGGCTGCGATCAGCATTTGAAGAAACGGTGAAGGCTTG 789
QY 832 CAGTTGCCGTATCAACCACTGATGGGAATGGGCTTTTGAAGAAACATTCCTT 891
DB 790 AATCTGCCGTGTGCTCATTTGATGGGCTGGGGCGTTTCCGGCAACGATCGTCA 849
QY 892 TCGGTGGATATGTTGGGTATGATGCGCACCGCTATGCCAATTTGCCGTCAGCAATG 951
DB 850 GCACTGGGCTATGCTGGGAATGACGCTGATCCTAGAAACCAATATGACATGCTAAC 909
QY 952 GATTTGTTATGCAAGTGGGGGCGGCTTTCAGACAGCCGGTAACTGGCAACTGAAGA 1011
DB 910 GATGTGATTTTCGCGGTGGGCTGACATTTTGAAGACGGAACGAACATCTGGCAAG 969
QY 1012 TTGTCTAGCCGCGCAAGTAATTCATTTGATGACATGACCCGCGGAGGTGGAAAAAC 1071
DB 970 TACTGCCCAATATCCACTGTTCTGCAATTCGATATGATTCCTTCAATTTCTAAAC 1029
QY 1072 AGGCTCCCGATGTGCCATTTGTGGGGATGTAAGCCATGTTTGAAGACGTTTTCAG 1131
DB 1030 GTACTGCGGATATCCGATGTTGGGGATGCTGCCAGGTCTCGAACAATGCTTGA 1089
QY 1132 CCGGCGCGGATATGGAATTAACCCACCATCCCATACCAACCCAGGCAATGTTAA 1191
DB 1090 CTCTTGTCCAAATATCCGCGCATCAACCACTGGAAGATCCGCACTGTGTGGCAGA 1149
QY 1192 ATTGATCATTTGCGGACCGCATTTACCCCTCCAGGTGCCCACTATGAGATCTATTG 1251
DB 1150 ATTGAACAGTGGGCGCTGTGTCAGTGTGCTGAATATGACATCACTGTAAGATTTAA 1209
QY 1252 CCCAGAGAGTATGACAGAAATGCTGCGCAGGCC---CGATGCTTACTACACACC 1308
DB 1210 CCGAGCGGTGATTCAGACTCTTTGCGGTTGACCAAGGAGACCTTACGATGAGTCC 1269
QY 1309 GATGTGGGCAACACCAAAATGTGGCGGCC---AGTTTGAACAATGAGCCCGCCGCA 1365
DB 1270 GATGTGGGCAACCAAGATGTTGCTGCACTTTATATTCATTCAGCAAAACCGGTCC 1329
QY 1366 TGGATTTCAATGCTGCTTGGGTAGATGAGGCTTTGTTTACCTGCGCCATGGAGCC 1425
DB 1330 TGGATCAATTCGCTGCTCGGCAAGATGAGGTTTGTTCGTCGCGACCTGGCGTCC 1389
QY 1426 AAGTGGAGTGGGGAGAGAG---CGGCAATTTGATAGTGGAGTGGCGCTTCCAAAG 1484
DB 1390 AATATGCGTTCGCAAGAAACCGGTGTTGGTATCTGGGAGCGCATTTCAAGATG 1449
QY 1485 AATCTTCAGAACTGGAAACCTTAGCCAGTACGACATCCAGTTAAACTTATTTCTC 1544
DB 1450 AATCTTCAGAACTGCTACCGGTTGCAATACGATTCGCCGTACTGTGTGATATCTC 1509
QY 1545 AATTAAGTGTGGAGGAGTGTGCTCAGTGGCAACAACTTTCTAGCAAGAGTAT 1604
DB 1510 AATTAAGTGTATCTGGGATGTGAGAGAGTGGCAGGACATGATTAATTCGGCGTCA 1569
QY 1605 TCTGCTTCAAGATG 1619

Db 1570 TCACATCTTATATG 1584

RESULT 12

US-10-282-122A-20222

Sequence 20222, Application US/10282122A

Publication No. US20040029129A1

GENERAL INFORMATION:

APPLICANT: Wang, Liangsu

APPLICANT: Zamudio, Carlos

APPLICANT: Malone, Cheryl

APPLICANT: Haselbeck, Robert

APPLICANT: Ohlsen, Kari

APPLICANT: Zyskind, Judith

APPLICANT: Wall, Daniel

APPLICANT: Trawick, John

APPLICANT: Carr, Grant

APPLICANT: Yamamoto, Robert

APPLICANT: Forsyth, R.

APPLICANT: Xu, H.

TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

FILE REFERENCE: ELITRA.034A

CURRENT APPLICATION NUMBER: US/10/282,122A

CURRENT FILING DATE: 2003-02-20

PRIOR APPLICATION NUMBER: 60/191,078

PRIOR FILING DATE: 2000-03-21

PRIOR APPLICATION NUMBER: 60/206,848

PRIOR FILING DATE: 2000-05-23

PRIOR APPLICATION NUMBER: 60/207,727

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: 60/230,335

PRIOR FILING DATE: 2000-09-06

PRIOR APPLICATION NUMBER: 60/230,347

PRIOR FILING DATE: 2000-09-09

PRIOR APPLICATION NUMBER: 60/242,578

PRIOR FILING DATE: 2000-10-23

PRIOR APPLICATION NUMBER: 60/253,625

PRIOR FILING DATE: 2000-11-27

PRIOR APPLICATION NUMBER: 60/257,931

PRIOR FILING DATE: 2000-12-22

PRIOR APPLICATION NUMBER: 60/267,636

PRIOR FILING DATE: 2001-02-09

PRIOR APPLICATION NUMBER: 60/269,308

PRIOR FILING DATE: 2001-02-16

Remaining Prior Application data removed - See File Wrapper or PALM.

NUMBER OF SEQ ID NOS: 78614

SOFTWARE: PatentIn version 3.1

SEQ ID NO 20222

LENGTH: 1815

TYPE: DNA

ORGANISM: Escherichia coli

US-10-282-122A-20222

Query Match 18.8%; Score 368.6; DB 13; Length 1815;

Best Local Similarity 55.3%; Pred. No. 2,1e-112;

Matches 805; Conservative 0; Mismatches 634; Indels 16; Gaps 4;

QY 172 CATGGGCTCAACACATTTTGGCTATCCGGCGGGCAATTTGGCCATCTATGAGAA 231

Db 139 CAGGCGTTAAACAAGTATCGATTATCCGGAAGCGAGTCTTGATATTATGATGCA 198

QY 232 CTGTACCGCTTGAAGCGCGGGGAAATTGAGCATATTTGTGCGCCATGAACAGAA 291

Db 199 -----TTGCAATACGTGGGTGATATGATCATGATATGATGTCATGAGCAAGCGG 249

QY 292 GCTTCCCATGGCGGAGATGGGTATCCAGACCCAGGTAAAGTGGAGTTTGTTCGT 351

Db 250 GCGGTGATATGGCGGATGGCTGCGCGCGGAGACCGGGAAGTGGCGTGTCTGCTGA 309

QY 352 ACATCTGAGACGAGGCGCACTAATTGGTGAACCGGCAATGGCCATTTGGACTCG 411

Db 310 ACGTGGGTCCAGGGGCGACCAATGCCATTCTGGCATCGCCACCGCTTATATGATTC 369

QY 412 GTGCCATGATGATGATTAATGAGAGGTGGGCGGTGCATGATTTGATGACCTTTC 471

Db 370 ATTCCATTAGTGTCTCTTCCGGGAGGTATGACACTGTTGATATGTTAGATGCTTT 429

QY 472 CAGAAATTAACATTTTGGCATCACTTACCATGTTTAAAGCACTCTTATGTGTACTG 531

Db 430 CAGGAGTGGACATGATGGGATTTTCGGACCGGTTTAAACAGATTTCTGTGTAAG 489

QY 532 AGTGGCGGATATGATGCTGCATGTTTCTGAGGCTTCCATCTTGTAGCACCGGTCTG 591

Db 490 CAACCGGAAGACATTTCCGAGGTGCTGAAAAAGGCTTTCTGTGGCGGCAAGTGTGC 549

QY 592 CCCGGGCGGTTTGTATGATATTCGAAAGATGTGGCTTGAAGAAATGTGAGTCAAT 651

Db 550 CCAGACCAAGTATGATGATTTACGAAAGATATCTTAATCCGGCCAGAAATTAACC 609

QY 652 CCCCTGACCCCGGAGATTAATCTACCGGTTATCCGCCACGTTAAAGTAAATGCC 711

Db 610 TATGCTGGCGGATGCTGATGCTTCAATCCCATCTACTACCGGACATTA 669

QY 712 CGACAAATTAATCGGCATTTGCAATTTTGAAGCAGGCGAGAAATCCCTTCTTACGTA 771

Db 670 GGGCAATTTAAGCGTCTCTGCAACGCTGGTAGCGGCAAAAAACCGTTGTCTACGTA 729

QY 772 GGGGAGGGGCGATGCGCCCAATGCCATGCCAGGTGCAAGAAATTTGGGAAAGTTC 831

Db 730 GCGGTGGGCAATACGCGCGGCTGCATAGCATTTGAAGAAACCGTGAAGCGTGTG 789

QY 832 CAGTCCGCTAACACACACCTGATGGGAATGGGCTTTTGAAGAAACCATTCCTCT 891

Db 790 AATCTCCCGTGTGTGCTCATATGATGGGGCTGGGGGCTTCCGGACACGATGCTAG 849

QY 892 TCGGTGGATATGTTGGATATGATGCAACGCTATGCAACTTTGCCGTACGGAATGT 951

Db 850 GCACGTGGGATGCTGGGAATGACCGGTACTACGAAGCAATATGACATCATTAACCG 909

QY 952 GATTTGTTGATGCACTGGGGGCGCGTTTCGACGACCGGATTAATGCAAACTAGACGA 1011

Db 910 GATGATATTTTGGCGTGGGATGATTTATGACCAAGCAAGCAAACTGTGCAAG 969

QY 1012 TTTGTAGCGCGCCAAAGTATTCATTTGATGACATGACCCGCGGAGGTGGAAAAAC 1071

Db 970 TACTGCCAAATGCACATGTTCTGATTCATATGATTCATCTTCATTTCTAAAC 1029

QY 1072 AGGCTCCGATGTCGATGTTGGGGAGATGAGGSCATGTTTGAACAGCTTTTGACG 1131

Db 1030 GTGATCGGATATCCGATGTTGGGGAGTGTCCAGGTCTGAAACAAATGCTTGA 1089

QY 1132 CGGCGCCGGGAATGATTTACCCACCATCCCATACACCCAGCATGTTAATGCG 1191

Db 1090 CTCTGTGCAAGATTCGCGCCATCAACCATGATGATCCGCACTGGTGGAGCAA 1149

QY 1192 ATTGATCATTTGGCGGACGATTAACCCCTCCAGGTGCCCATATGAGAGATATATGCC 1251

Db 1150 ATTGAACAGTGGCGGCTGCTGTCAGTCTGAATATGACATCAAGTGAAGATTAA 1209

QY 1252 CCCAGAGGTAGTACAGAAATTTGTGCGCAGGCC---CCGATGCTACTATCACACCC 1308

Db 1210 CCGGAGCGGTATGAGACTCTTTGGCGGTTGAGAGAGGAGACCGTTAGTGCATC 1269

QY 1309 GATGTGGCAACACCAATATGTGGCGGCC---AGTTTGAACATATGCCCCCGCCGA 1365

Db 1270 GATGTGGGACACACAGATGTTGCTGCACTTATTTATCAATTTGCAAAACCGGTGCG 1329

QY 1366 TGGATTTCCAGTGTGCTGGGTTCAGATGGGCTTTGTTTACCTGCGCCATGGAGCC 1425

Db 1330 TGGATCATTTCCGATGCTGCGGCTGCGGACGATGGTTTGTATCTGCGGCACTGGGCGT 1389

QY 1426 AAAGTGGAGTGGGGAGAG---CGGTCAATTTGATCAATGATGCAAGTCCAAATG 1484

Db 1390 AAATATGCGTTCCAGAGAAACCGTGTGTTGCGTCACTGGCGAGCGGAGATTTCAAGATG 1449

QY 1485 AATCTTCAGGAACGTGGAACCCCTAGCCAGTACGATCCAGCTTAAACTATTTATCTC 1544
 DB 1450 AACATTCAGGAACGTCTACCGCGCTTGCAATACGAGTTGCCCGTACGTGGTGAATCTC 1509
 QY 1545 AATAACGCTTGGCAGAGGATGCTGCGTACATGGCAACAACCTTCTACGAAGAAGCTTAT 1604
 DB 1510 AATAACGCTATCTGGGGATGCTGGAAGACATGGGACGATATCTATTCGCGCGCTGAT 1569
 QY 1605 TCTGCTTCTTAACATG 1619
 DB 1570 TCACAATCTTATATG 1584

RESULT 13
 US-10-282-122A-26894
 ; Sequence 26894, Application US/10282122A
 ; Publication No. US20040029129A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, Liangsu
 ; APPLICANT: Zamudio, Carlos
 ; APPLICANT: Malone, Cheryl
 ; APPLICANT: Haselbeck, Robert
 ; APPLICANT: Ohlsen, Kari
 ; APPLICANT: Zyskind, Judith
 ; APPLICANT: Wall, Daniel
 ; APPLICANT: Trawick, John
 ; APPLICANT: Carr, Grant
 ; APPLICANT: Yamamoto, Robert
 ; APPLICANT: Forsyth, R.
 ; APPLICANT: Xu, H.
 ; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
 ; FILE REFERENCE: ELITRA.034A
 ; CURRENT APPLICATION NUMBER: US/10/282,122A
 ; PRIOR FILING DATE: 2003-02-20
 ; PRIOR APPLICATION NUMBER: 60/191,078
 ; PRIOR FILING DATE: 2000-03-21
 ; PRIOR APPLICATION NUMBER: 60/206,848
 ; PRIOR FILING DATE: 2000-05-23
 ; PRIOR APPLICATION NUMBER: 60/207,727
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: 60/230,335
 ; PRIOR FILING DATE: 2000-09-06
 ; PRIOR APPLICATION NUMBER: 60/230,347
 ; PRIOR FILING DATE: 2000-09-09
 ; PRIOR APPLICATION NUMBER: 60/242,578
 ; PRIOR FILING DATE: 2000-10-23
 ; PRIOR APPLICATION NUMBER: 60/253,625
 ; PRIOR FILING DATE: 2000-11-27
 ; PRIOR APPLICATION NUMBER: 60/257,931
 ; PRIOR FILING DATE: 2000-12-22
 ; PRIOR APPLICATION NUMBER: 60/267,636
 ; PRIOR FILING DATE: 2001-02-09
 ; PRIOR APPLICATION NUMBER: 60/269,308
 ; PRIOR FILING DATE: 2001-02-16
 ; Remaining Prior Application data removed - See file Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 78614
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 26894
 ; LENGTH: 1833
 ; TYPE: DNA
 ; ORGANISM: Moraxella catarrhalis
 US-10-282-122A-26894

Query Match 19.8%; Score 368.2; DB 13; Length 1833;
 Best Local Similarity 53.9%; Freq. No. 2.8e-112;
 Matches 908; Conservative 0; Mismatches 743; Indels 34; Gaps 6;

QY 137 CCGGGGCTTATATCTGTAGTAGCTGAAGCGCATGGGCTCAACACATTTTGGCT 196
 DB 119 CAGGGCTGAGAGTGTCTTTCAGCATGTGATGAGAGTGTGAGATATTTTGGCT 178
 QY 197 ATCCGGGGGGAATTTTGGCATCTATGATGAATCTGTACCGCTTGAAGCGGGGG 256

DB 179 ATCCAGCGGTGGGATTTACATATTTATGATCTTTATTTAAGCAAGATA----- 229
 QY 257 AAATTGACATATTTTGGTGGCCATGAACAAGAGCTTCCATGGCGGATGGGTATG 316
 DB 230 AAATTGACATATTTTGGTGGCCATGAACAAGAGCTTCCATGGCGGATGGGTATG 289
 QY 317 CAGAGCAGAGGTAAGTGGAGTTTGTTCGTATCTGTGATCCAGAGGCGCATTAAT 376
 DB 290 GTGTAACACAGGACGAGCGGCTTTTGGCATTCAGGTCCAGGTGCAACCAACA 349
 QY 377 TGGTACCGGATCTGCAATGCGCATTTGGATCGGTGCGCATGTGTATCTAGAG 436
 DB 350 CTGTAAACAGCATTTGGAACCTGTATATGATTCATCCCAATGGTGTGTGGCAGTTC 409
 QY 437 AGGTGGCGGTGCGCATGATTTGGTATGATGATGATGATGATGATGATGATGATG 496
 DB 410 AGGTACCATCTAATTTGATTTGATGATGATGATGATGATGATGATGATGATG 469
 QY 497 CTTACCGATCGTTAAGCACTCCTATGTTAGTATGATGATGATGATGATGATGATG 556
 DB 470 CAGGCCCATCGTTAAGCACTCCTATGTTAGTATGATGATGATGATGATGATGATG 529
 QY 557 TTAAGAGCTTTCATCTTTGCTAGACCGGTGTCGCGGCGGCTTTGATGATGATG 616
 DB 530 TCAAAAAGCATTTTATATGTCAGGCTCGGCTGCTGCTGCTGCTGCTGCTGCTG 589
 QY 617 CCAAGATGTTGGGCTTAAAGATGATGATGATGATGATGATGATGATGATGATG 676
 DB 590 CAAGAATATGACCAACCAAGTATGATGATGATGATGATGATGATGATGATGATG 649
 QY 677 TACCGGATATCGCCCGAGGTTAAGTATGATGATGATGATGATGATGATGATGATG 736
 DB 650 TCAGATCATATGAGCTTACCAATTAAGTATGATGATGATGATGATGATGATGATG 709
 QY 737 TGTGAGACAGGCGCAAAATCCCTTCTCTAGTATGAGGAGGAGGAGGAGGAGG 796
 DB 710 CGCTACTATACGCAAAAGCGCTGATCTGATGATGATGATGATGATGATGATGATG 769
 QY 797 CCATGCCAGGTGCGAATTTGCGAAGGTTGCAATGTCGCGGTAACCAACCTGTA 856
 DB 770 CCAATTAAGAGCTGCGTCAATTTAGCGATGCTCAACTACTGTTACTAATACCTTA 829
 QY 857 TGGAAATTGGGCTTTTGAAGAAACCATCCCTTTCGTTGATGATGATGATGATG 916
 DB 830 TGGAGCTGGGTGCTTATCCAGGCTCAGGCAACAATTTGGAATGCTTGGATGATG 889
 QY 917 GCCACGGCTATGCGCACTTTGCGGTCAGGAGATGATGATGATGATGATGATGATG 976
 DB 890 GCAATATAGAGCAATGATGATGATGATGATGATGATGATGATGATGATGATG 949
 QY 977 GTTTCAGCAACCGGGTAACTGCGAATTAAGCAATTTTCTAGCCCGCAAGTAATTC 1036
 DB 950 GTTTGATGACGCTGAGCAATTAAGTCAAAAATTTGCGCAATGCAATGATGATTC 1009
 QY 1037 ACATGATACGACCGCGGAGAGTGGGAAAAACAGGGCTCCCATGAGGCCATTTGTC 1096
 DB 1010 AATGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1069
 QY 1097 GGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1156
 DB 1070 GCGATGTTAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1129
 QY 1157 CCGATCCCATACACCGCATGCTTAAATGCAATTTGATGATGATGATGATGATGATG 1208
 DB 1130 TCGATTAACATGCGCTCAATGATGATGATGATGATGATGATGATGATGATGATG 1189
 QY 1209 -----CGATTACCCCTCCAGGTGCGGATGATGATGATGATGATGATGATGATG 1261
 DB 1190 GTCTTGGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1249
 QY 1262 TAAATACAGAAATTTGTCGCGAGG---CCCGATGCTTACTACACCGATGATGATG 1318
 DB 1250 TGGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1309

QY 1319 AACACCAATGTGGGCGGCGCCAGTTT---TGAAACAATGGCCCCCGCCATGATTTCCA 1375
DB 1310 AGCATCAGATGTTTGGCTGCAATTTATATTAAGATGAGCCAGCCAAATGGTTAACT 1369
QY 1376 GTGCTGCTGGTACATGAGTGGCTTTGTTTACCTGCGCCATGGAGCCAAAGTGGAG 1435
DB 1370 CAGGTGGCTTGTAGTACATGAGCGGTGGTTTACTTATGCTATGAGCAAGAAATTTAGCTC 1429
QY 1436 TGGGGAGCAG-CCGTCAATTTGCAATGAGTGAAGTGGCAGCTTCCAAATGATTTAGG 1494
DB 1430 ACCCAAAAAGACAGTTTGTATGACTGGCGAAGGCTCATCCAAATGAATCCAG 1489
QY 1495 AACTGGGAACCTTACCCAGTACAGACATCCAGTTAAACTATTTCTCAATTAACGTT 1554
DB 1490 AGCTGTCACTTGTCTACAGTATATCTACAGTTAAATTTTAAATTAATATGCTC 1549
QY 1555 GCGCAGGGATGTGGGTGATGAGCAACAATTCTTCTAGAAAGCTTATTTGCTTTCTA 1614
DB 1550 AGCTTGGCATGTATTAAGAGTGGCAGATATGCTTTATGAAAGTGGCAGTCAATCTT 1609
QY 1615 ACATGTCCAGGAGTGCAGACATTAATCTCTGTGAAAGCTATGAGCATCAAGGTA 1674
DB 1610 ATATG---CAATCTTTACTATTTTGTAAATTTGAGCAAGATTAAGGCAATTAAGTG 1666
QY 1675 TTACTGTGCGCAGCGGAGATTTGGCCCGCGGATCGCCGAATGCTAGCCCAATG 1734
DB 1667 TAAATAATTAACAATCTGCTACATGACAGAGAGCTAAAAAAGCCCTAGAAATGATG 1726
QY 1735 GTCCGTGTGATGATGATGTGTGTGCAAAAAGATGAAACCTGTACCTATGATTTGCC 1794
DB 1727 GTTGTGATTTTATGATGTCTATATGTGATTAAGGTGAACATGTTATCCAGATGACGTG 1786
QY 1795 CCGGC 1799
DB 1787 CTGGC 1791

RESULT 14
US-10-672-787-35
; Sequence 35, Application US/10672787
; Publication No. US20040067554A1
; GENERAL INFORMATION:
; APPLICANT: LAGACE, Robert, E.
; APPLICANT: BERG, Kim, L.
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF MORAXELLA CATARRHALIS GENOME
; FILE REFERENCE: ELITRA.025CI
; CURRENT APPLICATION NUMBER: US/10/672,787
; CURRENT FILING DATE: 2003-09-26
; PRIOR APPLICATION NUMBER: 09/596,002
; PRIOR FILING DATE: 2000-06-16
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PERL Program
; SEQ ID NO 35
; LENGTH: 96109
; TYPE: DNA
; ORGANISM: Moraxella catarrhalis
US-10-672-787-35

Query Match 18.8%; Score 368.2; DB 13; Length 96109;
Best Local Similarity 53.9%; Pred.No. 3.1e-111;
Matches 908; Conservative 0; Mismatches 743; Indels 34; Gaps 6;

QY 137 CCGGGGCTTATATCTGATGATAGCCGGAACCCGCTGAGCAACATTTTGGCT 196
DB 2433 CAGGGCTGAGATGCTTGTCAAGCATGATGATGAAGGTGAGATATTTTGGCT 2492
QY 197 ATCCGCGCGGCAATTTTGGCCATCTATGATGAAGTACCGCTTTGAAGCGCGGGG 256
DB 2493 ATCCGCGCGGCGGTATTAATATTTATGATGCTTTATTAAGCAAGATA----- 2543
QY 257 AATTTGACATATTTTGTGTGCGCCATGAACAAGAGCTTCCCATGCGCGGATGGTATG 316

DB 2544 AATTTGACATATTTTGTGTGCGCCATGAACAAGCGCGGCAATATGGCCGATGCATCA 2603
QY 317 CCAGAGCCACAGGTAAAGTGGAGTTTGTTCGTAATCTGACACAGGGGCGATCACT 376
DB 2604 GTGTACAAACAGCGAGAGCGGCGTTTGTGGGCACTTCAGGTCCAGGTCAACCA 2663
QY 377 TGGTACCGGATTCGCAATGCCATTTGGAAGTGGTCCCATGATGGTGAATTAAGAG 436
DB 2664 CTGTAAACCACTTCCACTGCTTATATGATTCATATCCCATGATGTTGTGTGCAAGT 2723
QY 437 AGGTGGCGGCGCATGATTTGTAAGCATGCTTTCCAGAAATTTGACATTTTGGCATCA 496
DB 2724 AGGTACCAATTAATTTGATGATGATGATGATGATGATGATGATGATGATGATG 2783
QY 497 CCTACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 556
DB 2784 CAGCGCCATGCTTATAGCAAGTTTCAAGTTGCGCAGCGGAAGACATTTCCAAATCA 2843
QY 557 TTACTAGGCTTTTCATCTTGTAGACACCGGTGTCGCGGCGGTTTGAATGATATTC 616
DB 2844 TCAAAAAGCATTTTATTTATGAGGCTGCGGTCGTCGTCGTCGTCGTCGTCGTCG 2903
QY 617 CCAAGATGTGGGCTTAAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 676
DB 2904 CAAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2963
QY 677 TACCGGATATGCCCCAGTAAAGTAAATCCCGCAATTAATGCGCATTTGCAAT 736
DB 2964 TCAGATCAATCAGCTTACATTAAGGTCAATGCGGCAATCAAAAAGCCATCGAAA 3023
QY 737 TGTGTAGACAGGCGGAATCCCTGCTTACGTATGAGGAGGAGGCGCATTCGCCCATG 796
DB 3024 CGCTACTATCAGCCAAAGCGCTGTACTGTATTCAGGCGGTGGGTGGTGGTATG 3083
QY 797 CCGATGCCAGGTGAGGATTTTGGGAAAGTTCCAGTTGCGGTAACAACCCCTGA 856
DB 3084 CCAGTAAAGAGGTGCTCATTTAGCCGATAGCCGATTAACCTATACCTTAA 3143
QY 857 TGGGAATTTGGGCTTTTGAACAAAACATCCCTTTGCTGAGGATGATGATGATGATG 916
DB 3144 TGGGACTGGGTCTTATCAGGCTCAGCGAACAATTTTGTGATGCTTGTATGATG 3203
QY 917 GCCACGGGTATGCCAATTTTGGCCGTCAGCATGATGATTTGTAATGCAATGGGAGCC 976
DB 3204 GCACATATGAGCAACATGACATGATGATGATGATGATGATGATGATGATGATG 3263
QY 977 GTTTCAGCAGCGGTTAATCTGCAACATGACGAAATTTGCTAGCGCGCCAAATGATTC 1036
DB 3264 GTTTGATGACCGGTGACCAATATGCAAAAATTTGCGCAATGCAATGATGATTC 3323
QY 1037 AATTTGACATGACCGCGGAGAGTGGGAAAAAAGGCGTCCGATGTCGCAATTTG 1096
DB 3324 ATATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3383
QY 1097 GGGATGATGCGCATTTTGAACAGCTTTTGCAGCGGCGCGGAAATTTGATTAACCCA 1156
DB 3384 GCGATGTTAAGTGTGATCAATTAATGATGATGATGATGATGATGATGATGATGAT 3443
QY 1157 CCCATCCCATACACCCAGGATGTTAAATGATGATGATGATGATGATGATGATGATG 1208
DB 3444 TGGATGAATGCGCTCATATGTTGTGCAACAATTAATGAATGGGCTTAAAGCCATG 3503
QY 1209 -----CGATTAACCCCTCCAGGTGCCCATATGAGATATCAATGCCCCCAGAGG 1261
DB 3504 GTCTTGCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3563
QY 1262 TAGTACAGAAATTTGATGCGCAGG---CCCGATGCTTACTACACACCGATGAGGAG 1318
DB 3564 TGGTGAAGTGTCTTATATAGCTGAAGATGATGATGATGATGATGATGATGATGATG 3623
QY 1319 AACACCAATGTGGGCGGCGCCAGTTT---TGAAACAATGGCCCCCGCCATGATTTCCA 1375

Db	3624	AGCATACAGATGTTTCCTGCATGTGTTATTAAGTATGATGAGCAGCAAGTTAACT	3683
Qy	1376	GTGCTGGGTTTGGGATCCGATGGGCTTTGGTTTACCTGGCCCATGGAGCCAAAGTGGAG	1435
Db	3664	CAGGTGGCTTAGGTAACAATGGGCGCTTGGTTTACCTTATGCTATGGCAACAAATTAAGCTC	3743
Qy	1436	TGGGGAGCAG- CGGTCAATTTGCATCAGTGGAGATGCAGCGCTCCAAATGAATCTTCAGG	1494
Db	3744	ACCCAAAAAAGCAGTTGTTTGTATCACTGGGAGAGGCTCAATCAATGAATCCAG	3803
Qy	1485	AACTGGAAACCTAGCCCAATGACATCCAGGTTAAACTATTAATCTCAATAACGTT	1554
Db	3804	AGCTGTCAACTTGTCTACAGTAAATCTAACAGTTAAATTTAAATTAATATGCTC	3863
Qy	1555	GCGAGGGAGTGGTGGCGTACGTGGCAACAACTTTACAGAGAACGTTATCTGTTCTA	1614
Db	3924	ATATG---CAATCTTTACTGTATTTGTTAAATTTGGCAGAAAGTATGGCATTAAGGTG	3980
Qy	1675	TTACTGTCCGAAAGCGGGAAGATTTGGCCCCGGCGATGCCGAAATGCTAGCCACAATG	1734
Db	3961	TAAAAATTAACCAATCTCTCTACCAAGCAAGAAAGACTTAAACAGCCCTTAATAATGATG	4040
Qy	1735	GTCCGTGGTATGATGATGTGGTGGTCAAAAAAGATGAACACTGTTATCCCTATGATTGCC	1794
Db	4041	GTTTGGATTATTATGATGATGATGATGATTAAGGTAACAGTATTCGATGCAAGTGC	4100
Qy	1795	CCGCG 1799	
Db	4101	CTGGC 4105	
RESULT 15			
US-10-282-122A-30934			
; Sequence 30934, Application US/10282122A			
; Publication No. US20040029129A1			
; GENERAL INFORMATION:			
; APPLICANT: Wang, Liangsu			
; APPLICANT: Zamudio, Carlos			
; APPLICANT: Malone, Cheryl			
; APPLICANT: Haeselbeck, Robert			
; APPLICANT: Ohlsen, Karl			
; APPLICANT: Zyskind, Judith			
; APPLICANT: Wall, Daniel			
; APPLICANT: Trawick, John			
; APPLICANT: Carr, Grant			
; APPLICANT: Yamamoto, Robert			
; APPLICANT: Forsyth, R.			
; APPLICANT: Xu, H.			
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms			
; FILE REFERENCE: ELTPA.034A			
; CURRENT APPLICATION NUMBER: US/10/282,122A			
; CURRENT FILING DATE: 2003-02-20			
; PRIOR APPLICATION NUMBER: 60/191,078			
; PRIOR FILING DATE: 2000-03-21			
; PRIOR APPLICATION NUMBER: 60/206,848			
; PRIOR FILING DATE: 2000-05-23			
; PRIOR APPLICATION NUMBER: 60/207,727			
; PRIOR FILING DATE: 2000-05-26			
; PRIOR APPLICATION NUMBER: 60/230,335			
; PRIOR FILING DATE: 2000-09-06			
; PRIOR APPLICATION NUMBER: 60/230,347			
; PRIOR FILING DATE: 2000-09-09			
; PRIOR APPLICATION NUMBER: 60/242,578			
; PRIOR FILING DATE: 2000-10-23			
; PRIOR APPLICATION NUMBER: 60/253,625			
; PRIOR FILING DATE: 2000-11-27			
; PRIOR APPLICATION NUMBER: 60/257,931			
; PRIOR FILING DATE: 2000-12-22			
; PRIOR APPLICATION NUMBER: 60/267,636			

```

; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO: 30934
; LENGTH: 1725
; TYPE: DNA
; ORGANISM: Pasteurella multocida
US-10-282-122A-30934

Query Match      18.8%; Score 367.8; DB 13; Length 1725;
Best Local Similarity 54.8%; Pred. No. 3,76-112;
Matches 821; Conservative 0; Mismatches 662; Indels 16; Gaps 4;

QY      128 AACGGCAAAACCGGGCTTATATCTGTATGATAGTACCTGAAACGCCATGAGGGTCAACACA 187
Db       5 AAAAGTATCTGCGCGCAAAAATGCTGTGTGACGTCTTTCGCGATGAAAGCGCTGAGTTT    64

QY      188 TTTTGGCTATCCCGGGGGGGGSCAATTTTGGCCATCTATGATGAACTGATACCGCTTTGAG 247
Db       65 TATTGGTTATCCCGGGGATCGGTTTTGATATTATATGATGCAATTCATACCTT----- 119

QY      248 CGGGGGGGGAATAGACATATTTTGGTGGCCATGAAACAAGAGACTTCCCATCGGCGG 307
Db      120 ----GGGGGGGATTGAACATATCTTGTTCGACATGAAACAAGCGCTGTGACATGGAG 175

QY      308 ATGGGTATGCCAAGGCCACAGGTAAATGGGAGTTTGTTCGTAACATCTGGACCAAGGG 367
Db      176 ATGGTATGCGCAAGGCCACCGTAAATGGGTGTGTGTTCTTGTAACAATCCGAGCCGGTG 235

QY      368 CGACTAATCTGGTACCGGCAATTTGCCAATGCCCATTTGACCTCGTGGCCATGGNGTGA 427
Db      236 CAATTAATACCATCACTGTGATTTGCCAACAGGATATAGATTCCTGCTTGTATTT 295

QY      428 TTACTGAGAGGTGGGCGGTGCACATGATTGGTAGCATGCTTTCCAGAAATTGACATTT 487
Db      296 TATCTGACAGAGTGATGTCAAGATTAATTGGACATGATGCTTTCAAGAAATGCATATGG 355

QY      488 TTGGCATCACTTAACGATGGTTAAGACACTCCATGATGATACGTAAGTGGCGCGGATATNG 547
Db      356 TGGTATCTCTCGTCCACGTGCTCAACAACAGTTTTTAGTGAATCTGGCGCTGAATTC 415

QY      548 CTCGCAATGTTACTGAGGCTTTCCATCTTGTGACACCGGTGTCGCCGGCGGCTTTTGA 607
Db      416 CTGAAACCATTAATAAAGCTTTTATATTTGTGGTCAACAGGAGACCGGACCTGTGGGA 475

QY      608 TCGATTAATCCCAAGAGATGGGCTTAAGAAGATGTAGTACATTCCTCTGACCCCGGTG 667
Db      476 TCGATTTGCCAAGAGATGGTAAATCTGCACATAAATTAATCAATTAATCAAGAA 535

QY      668 ACCTTAATCTACGGGGTTATCGCCCAACGGTTAAAGTAAATCCCCGACAAATTATCGG 727
Db      536 GAGTTGAATTACGTTCTTAATACCAACAAATPACAAGGCAATPAAAGGCAATTAATAAAG 595

QY      728 CATTGCAATGTTGAGCAAGCCCAAGAAATCCCTGCTCTACGTAGGAGGAGGCGCATCG 787
Db      596 CATTAATAAGCCTTAATTAATGTGCGCAAAAAGCGGTTATTGTGTGTGGTGTGTATCA 655

QY      788 CCGCCAAATGCCAATGCCAGATGCAAGAAATTTGGCGGAAAGGTTCCAGTTGCGGGTAAACA 847
Db      656 CTGGGAATTTGAGGCAAAATTAAACGAATTTTGACAAACCTTAAATTTGCGGTCACTT 715

QY      848 CCAACCTGATGGGAATTTGGGACTTTTGGACGAAAAACAATCCCTTTGCGTGGGATATGTGG 907
Db      716 CCTCCTTAATAGGGAATTTAGGTGCTATCAAGCTCAGATTAAGCAATTTCTTAGGGAGTTAG 775

QY      908 GTATGATATGGCCACCGGTATAGCAACTTTTGGCCGACAGCGAATGATGATTTGTGATTTGCG 967
Db      776 GCAATGCAATGTAATTGAAGAAAGCAATATATCAATGATCACTAATGATCTCAATTTTGTA 835

QY      968 TGGAGGCGCGTTTCAAGCAACGGGTAACTGGCAAACTGAACGAATTTTGTAACCGCGCCA 1027

```

Db 836 TTGGTGGCGCTTTTGATGATGAAAGCACTAATACTGGCAAAATATTGTCCACATGCAA 895
QY 1028 AAGTAATTACATGATGACATGACCGGGGAGTGGGAAAAACAGGGGCTCCGATGTGC 1087
Db 896 AAGTGAATTCATATTGATGATGATGATGATGATGATGATGATGATGATGATGATG 955
QY 1088 CCATTGTGGGGATGTACGCCATGTTTGAAGAGCTTTTGCAGCGGGCCCGGGAATTGG 1147
Db 956 CAATTGTGGGAGTGCAGAAAAAGTCAATTAATGATTTTAACTTATAGAGATGACA 1015
QY 1148 ATTACCCACCCATCCCATACCAACGAGCATGTTAAATCCGATGATCATTTGGCGGA 1207
Db 1016 ATTATTTAAATTCGAAAATGCTTGGATGATGATGATGATGATGATGATGATG 1075
QY 1208 CCGATTACCCCTCCAGGTGCCCACTATGAGATTAATGCCCCCAGAGATGATAC 1267
Db 1076 CAGTAAATGTCTCTCTTTTACGAGATGATTCGTCATTAACCAAGCAAGATGATGG 1135
QY 1268 ACGAAATGTGTGCCAGGCCCCCGAT--GCTTACTACACCAAGATGTGGACAAACCC 1324
Db 1136 AAGTGTATTATCGATTAAACAAAGGTAGGCTTATGTGCTCAGATGTTGACAAATC 1195
QY 1325 AAATGTGGGGCGGC--CCAGTTTGTAAACAATGAGCCCCCGCGATGATTTCCAGTGTG 1381
Db 1196 AAATGTGTGGGCAATTACATTAATCCGTTTAAACAACCTCGGCGTTGATTAATCTGTG 1255
QY 1382 GCTTGGGTACGATGGGCTTTGTTTACTGTGCCCATGGAGCCAAAGTGGAGTGGGG 1441
Db 1256 GCTTAGGGACCATGGGCTTTGACTACCGCCGCACTTGGGTAAAGCTGCTTATCTG 1315
QY 1442 ACG-AGCGGTCAATTGATCAGTGGAGATGCCAGCTTCAAATGAATCTTCAGGAACCTGG 1500
Db 1316 ACTCAACAGTGTGTGTGTGTGACGGGGATGGCAGTATTCAAATGAATATTCAGAGCTTT 1375
QY 1501 GAACCTTAGCCCACTAGACATCGAGTTAAACTATTAATTCGAATACGTTGGCAGG 1560
Db 1376 CTACTGCACAGCAATATGATATTCGAATGTGATTATCAGTTGAATACCGCTTTAG 1435
QY 1561 GGATGTGCGTCAATGGCAACAATCTTCTAGAAAGATGATTTGCTTCAAGATG 1619
Db 1436 GCAATGTGAAACAGTGGCAAGATCTGATTTATTTCTGAGCGCATTCAGATCTTATATG 1494

Search completed: July 25, 2004, 13:28:24
Job time : 952.797 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 25, 2004, 03:59:07 ; Search time 5589.05 Seconds

(without alignments)
10466.900 Million cell updates/sec

Title: US-09-893-033-6

Perfect score: 1959
Sequence: 1 gccataaggagcccatcgccg.....cataagcacaattgaattc 1959

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estdb:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_est3:*
12: gb_est4:*
13: gb_est5:*
14: gb_est6:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_hiv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rtd:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	265.4	13.5	2673	11	AY109387 Zea mays
2	204.4	10.4	860	14	CD446001 EL01T0206
3	181.6	9.3	890	29	CG219527 CG2BK677V
4	178.8	9.1	890	13	CA092873 SCCCC1100

Result No.	Score	Query Match	Length	DB ID	Description
5	176.8	9.0	826	28	BH390390 AG-ND-153
6	176.6	9.0	768	12	BG595301 EST4943979
7	175.6	9.0	775	12	BG599315 EST504215
8	175.2	8.9	811	29	CC722281 CG0EB62TV
9	174.6	8.9	1121	14	CK166070 FGAS05012
10	173.2	8.8	812	14	CB626704 OS1EB01C
11	170.8	8.7	897	29	CG970807 MBENF13TF
12	170.6	8.7	708	29	CG851756 ZMWB034
13	169	8.6	1036	29	CNS067BR
14	168.6	8.6	762	14	CD446176
15	164.6	8.4	802	12	B1935831 EST555720
16	159.6	8.1	1149	14	CK167517 FGAS05189
17	159	8.1	796	14	CF118966 MTU10CS.P
18	156.2	8.0	753	12	BM412862 EST587189
19	154.8	7.9	740	12	BG596565 EST495243
20	154.6	7.9	770	12	B1935305 EST555194
21	154.6	7.9	1143	14	CK166817 FGAS05104
22	153.8	7.9	737	13	BU099465 WHE3307.F
23	150.2	7.7	717	29	CG606643 CGDWG71TH
24	150.2	7.7	857	14	CD440317 EL01N0553
25	150.2	7.7	886	29	CC992548 ZUAAF53TH
26	150	7.7	725	13	BQ60365 QGC15J01.
27	150	7.7	726	13	BQ63809 QGC24013.
28	149.4	7.6	658	13	BQ85987 QGC14118.
29	149.4	7.6	683	13	BQ856535 QGC5103.Y
30	149.4	7.6	685	13	BQ863133 QGC23A23.
31	149.4	7.6	691	13	BQ863304 QGC23I09.
32	149.4	7.6	704	13	BQ860226 QGC15C23.
33	149.4	7.6	705	13	BQ861331 QGC18D02.
34	149.4	7.6	713	13	BQ864873 QGC27016.
35	149.4	7.6	729	13	BQ867096 QGC9W11.Y
36	149.4	7.6	749	13	BQ866590 QGC8610.Y
37	148.4	7.6	666	13	BQ863263 QGC23G16.
38	148.4	7.6	685	13	BQ866269 QGC7103.Y
39	148.4	7.6	689	13	BQ858743 QGC11B17.
40	147.8	7.5	608	14	CF441997 EST678342
41	147.8	7.5	665	13	BQ862151 QGC30E16.
42	147.8	7.5	687	13	BQ864317 QGC26G04.
43	147.8	7.5	702	13	BQ863370 QGC23L04.
44	147.8	7.5	734	13	BQ864734 QGC27107.
45	147.8	7.5	742	13	BQ861736 QGC19H05.

ALIGNMENTS

RESULT 1
AY109387
LOCUS 2673 bp mRNA linear HTC 17-OCT-2002
DEFINITION Zea mays CU2281_2 mRNA sequence.
ACCESSION AY109387
VERSION AY109387.1 GI:21213093
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Direct Submission
Submitted (25-APR-2002) Maize Mapping Project, University of Missouri, Columbia, MO 65211, USA
If you are interested in getting corresponding physical clones, these are publicly available in getting from ZmDB, www.zmdb.iastate.edu, TIGR, www.tigr.org, or NCBI, www.ncbi.nlm.nih.gov. When the source of

maize cDNA sequences is either Virginia Malbot, Stanford or Pat Schnable, Iowa State, then clones may be requested from ZmDB: www.zmdb.iastate.edu.

FEATURES
Location/Qualifiers

1..2673
/organism="Zea mays"
/mol_type="mRNA"
/db_xref="MaizeDB:630902"
/db_xref="taxon:4577"
/clone_lib="Maize Mapping Project/Dupont Consensus Library"
/note="this sequence is part of a project of EST assemblies resulting from the application of public contigs to seed Dupont contigs; this resource was assembled by Dupont as part of a collaboration for the overgo addressing of BACs in conjunction with the Maize Mapping Project"

ORIGIN

Query Match 13.5%; Score 265.4; DB 11; Length 2673;
Best Local Similarity 50.9%; Pred. No. 1,26-65;
Matches 757; Conservative 0; Mismatches 697; Indels 33; Gaps 8;

QY CCTAGCCACGCGAAACCGGGCTTATATCTGTAGTAGAGCTGAAGCCATGGGGTC 180
DB CCCACGAGCCCCCGCAGAGGGCGCGACATCTCTGTCAGTCCCTCGAGCCCTGGGGTC 638
QY AAAACATTTTGGCTATCCCGCGGGCAATTTGGCCATCTATGATGACTGTACCGC 240
DB CGCGACGCTTCGCTACCCCGCGGGCGCTGCATGAGATCCACCGGACATCACCCGC 758
QY TTTGAGAGCGGGGGAATAGACATATTTGTGGCCATGAGCAAGAGCTTCCCAT 300
DB TNN 759
QY GCGCGGATGGGTAATGCCAGACCAAGTAAAGTGGAGTGTGTTCCGTAATCTGA 360
DB GCCGCTCCGCTCAGCGCGCTCCTCGGCGCGCTCGCGCTCANNNNNNNNNNNNNN 869
QY CCAGGGGGAATCTTGTGATCGCGCATTTGCCAATGCCATTTGACTGGGCGCATG 420
DB NNN 870
QY GTGTGATTTCTGAGAGGTGGCGGCTGCAATGATGATGATGATGATGATGATGAT 480
DB GTGCGCATCAGGGAGAGGTGGCGGAGCATGATGATGATGATGATGATGATGATGAT 989
QY GACATTTTGGCATCCTTACCGATGCTTAAGCATCTTATGATGATGATGATGATGAT 540
DB CCCATGCTGAGAGTCAACCGCTCATCAACAGCAACAATCACTGATGCTGAGCTGAC 1049
QY GATAGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
DB GACATNN 1050
QY GTTTGATGATGATTTCCCAAGATGAGGCTTAAAGAAATGATGATGATGATGATGAT 660
DB GTGCTTGTGACATCCCAAGACATCC---ACAGAGATGAGGCTGCGGCTGGGA 1165
QY CCCGTAAGCTTAATCTACCGGGTATAGCCCAAGCTTAAGTAAATCCCAAGAT 720
DB CACGCCCTAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1225
QY AATGCGGAT---TGCAATTTGAGAGAGCGCAAGAAATCCCTGCTTACGATGAGGG 776
DB GCTTGAAGAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1285
QY AGGGGCGATCGCCGCAATGCCATGCCAGTGAAGAAATTCGCAAGATTCAGAT 836
DB TGG-----CTGCGAGCATCTGGAGAGATTCGCGGCTTGTGAGCTGACTGAAT 1339
QY GCCGTAACAACCCCTGATGAGAAATTTGGGCTTTTGAACGAAACCATCCCTTTCGT 896

DB 1340 CCCGTAACAACCCCTGATGAGAAATTTGGGCTTTTGAACGAAACCATCCCTTTCGT 1399
QY 897 GGGATGTTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 956
DB 1400 GCGCATGCTGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1459
QY 957 GTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1016
DB 1460 GTTGTTCACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1519
QY 1017 TAGCCGCGCAAGATTAATTCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1076
DB 1520 AGCAGGCGTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1579
QY 1077 TCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1136
DB 1580 GCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1636
QY 1137 CCGGAAATGATTAATCCCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1196
DB 1637 TCTTGAAGAGACATCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1696
QY 1197 TCATTGGCGGACCGATTAATCCCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1256
DB 1697 TCAGCAAGAGAGGATTAATCCCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1756
QY 1257 GAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1313
DB 1757 ATATCTATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1816
QY 1314 GAGCAACAACCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1370
DB 1817 TGGCAGACCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1876
QY 1371 TTCCAGTGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1430
DB 1877 GTTTCACCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1936
QY 1431 GGG-AGTGGGAGCAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1489
DB 1937 GGCACACCAAGTGTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1996
QY 1490 TAGGAATGGAACCTTAAGCCAGTACGATGATGATGATGATGATGATGATGATGATGATGATGAT 1549
DB 1997 TCAGAGCTAGCTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2056
QY 1550 CGGTTGGAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1596
DB 2057 CCAGCATCTGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2103

RESULT 2
CD446001 860 bp mRNA linear EST 03-JUN-2003
LOCUS E101T0206D04.b Endosperm_4 Zea mays cDNA, mRNA sequence.
DEFINITION
ACCESSION CD446001.1 GI:31361644
VERSION
KEYWORDS
SOURCE
ORGANISM
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
Clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 860)
REFERENCE
Lat,J., Dey,N., Kim,C.S., Becraft,P., Larkins,B., Linton,E. and
Messing,J.
Sequencing of the maize endosperm ESTs
Unpublished (2002)
TITLE
JOURNAL
CONTACT: Lai, Jinseng
DR. Joachim Messing's lab
Waksman Institute, Rutgers University
190 Frelinghuysen Rd., Piscataway, NJ 08854, USA
Tel: 732-445-3801

Fax: 732-445-5735
Email: jlai@waksman.rutgers.edu
Seq primer: 73.

FEATURES

Location/Qualifiers
1..860
/organism="Zea mays"
/mol_type="mRNA"
/cultivar="W22"
/db_xref="taxon:4577"
/clone_type="Endospem_4"
/clone_lib="Endospem_4"
/note="Vector: pbluescript SK-, Site_1: EcoRI; Site_2: XhoI"

ORIGIN

Query Match 10.4%; Score 204.4; DB 14; Length 860;
Best Local Similarity 54.0%; Pred. No. 5e-48;
Matches 443; Conservative 0; Mismatches 371; Indels 6; Gaps 1;

236 CCCATGCGCGGATGGTATGCGACGACGAAAGTGGAGTTTGTTCGGTACAT 355
11 CTTTGGCGGCTCCGCTACGCGCCTCTCGGGCGCGCTCGGCTCGCATCGCACCT 70

356 CTGACACAGGGGCGACTTAATTTGTGACCGGATTTGCCAATGCCATTGTGACTCGGTGC 415
71 CCGGCCCCGGCGCCACCACTAGTCTCGGCTGCGGACGCGTGTCTCATTTCTGCTCC 130

416 CCATGATGATTAATCTGAGAGAGTGGCCGCTGCCATGTTGTAGCGATCTTCCAGG 475
131 CCATGATGATTAATCTGAGAGAGTGGCCGCTGCCATGTTGTAGCGATCTTCCAGG 190

476 AAATTGACATTTTGGCATCACTTACCGATTCGTTAAGCATCTCTATGTGTAGTATGT 535
191 AGACGCCATGCTGAGGTCAACCCGCTCATACCAACACACAACTCACTGATCTCGACG 250

536 CGGCGGATATGCGCTGCAATTTGTAATGAGGCTTTCATCTGCTAGACACCGGATGTCGG 595
251 TCGACGACATCCCGCGGCTCGTGACGAGAGGCTTCTCTCGGCTCGCTGCTGTGACAG 310

596 GCGCGGTTTGTATGATATATCCCAAGATGTGGGCTTGAAGATGTGATACATTTCCC 655
311 GCGCGGTTTGTATGATATATCCCAAGATGTGGGCTTGAAGATGTGATACATTTCCC 370

656 TCGACCCCGGATGAGCTTAATCTACCGGCTTAATCGGCCACGAGTTAAAGTAAATCCCGAC 715
371 ACACACCATAGATGTGCTGCGGTACATTTGGCGGCTTCCAAAGCCCTTGGACATGAT 430

716 AAATTATGCGGATGCAATTTGTTGAGACAGGCGCAAAATCCCTGCTTACGTAAGGG 775
431 TCGTTGACAGAGTGTGCGTCTTGTGTTGTAATCCGCGGCTGCTTATATGAGGCG 490

776 GAGGGGCGATGCGCGCCCAATGCCATGCGGAGTGCAGAAATTTGCGGAAAGTTCAGT 835
491 GTGG-----CTGGCAGCATGTGTGAGAGAGTTGCGAGCTTGTGAGCTGAGAA 544

836 TCGCGGTAACAACAACCTGATGAGAAATGGGGCTTTTGAAGAAACCATCCCTTTGCG 895
545 TCGCGGTAACAACAACCTGATGAGAAATGGGGCTTTTGAAGAAACCATCCCTTTGCG 604

896 TGGGATATGAGGATATGATGAGCCACCGCTATGCCAATTTTGGCGTCAAGCAATGTGAT 955
605 TGGGATATGAGGATATGATGAGCCACCGCTATGCCAATTTTGGCGTCAAGCAATGTGAT 664

956 TGTATATGAGTGGGCGCGCTTTCAGACACCGGGTAACTGCGAAATGACCAATTTG 1015
665 TGTATATGAGTGGGCGCGCTTTCAGACACCGGGTAACTGCGAAATGACCAATTTG 724

1016 CTAGCCGCGCCAAATTAATTCATGATGACATGACCCGGCGAGGTGGGAAAAAACAAGG 1075
725 CAAGCAGGCGCTAAGATTGTGACGCTTGATATTGATCCGCTGAGATTGGCAAGAAACAAG 784

1076 CTCCGATGTGCCATTTGTTGGGGATGTAACGCATGTTT 1115

Db 785 AGCCACATGTGCCATCTGTGACAGATTTAAGCTTCTT 824

RESULT 3

CG219527/c 890 bp DNA linear GSS 22-AUG-2003
LOCUS CG219527/c 890 bp DNA linear GSS 22-AUG-2003
DEFINITION O22BK67TV_ZM_0.7.1.5_KB Zea mays genomic clone ZMMEM40753K13,
genomic survey sequence.
ACCESSION CG219527
VERSION CG219527.1 GI:34119415
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 890)
Whitehead, C.A., Quackenbush, J., Van Aken, S., Uterback, T.,
Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Konliff, T.,
Citek, R.W., Nunberg, A., Robbins, D. and LaKey, N.
Consortium for Maize Genomics
Unpublished (2002)
Other GSS: O22BK67TV
Contact: Cathy Whitehead
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitehead@tigr.org
Seq primer: TF
Class: sheared ends.

FEATURES

Location/Qualifiers
1..890
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone_lib="ZMMEM40753K13"
/note="Vector: pBCK-; Site_1: HincII; 0.7-1.5 Kb
methylation filtered genomic DNA library"

ORIGIN

Query Match 9.3%; Score 181.6; DB 29; Length 890;
Best Local Similarity 52.4%; Pred. No. 2.5e-41;
Matches 456; Conservative 0; Mismatches 399; Indels 15; Gaps 2;

121 CTTAGCCACGCGCAACCGGGCTTATATCTGATGATAGCTGAAACGCAATGGGCTC 180
857 CCCACGATCCCGCAAGGCGCGGACATCTGTCGAGTCCCTGAGCGCTGCGGCGTC 798

181 AAACAATTTTGGCTATCCCGGCGGCGCAATTTGCCATCTATATGAATCTTACCGC 240
797 CGGAGGCTTTCCTCAACCCCGGCGGCGTCCATGAGATCCACGACACTCAACCCG 738

241 TTTGAACGCGCGGGGAAATTGACATATTTTGTGCGGCAATGACAGAGCTTCCAT 300
737 T-----CCCCGATTCGCAACACACTTTCGCGCAGAGGAGGCGCTTT 687

301 GCGCGGATGGGTATATGCGACAGGCAAGTAAAGTGGAGTTTGTTCGATCATTTGGA 360
686 GCGGCTTCGCGCTACCGCGCTCTCGGCGCGGCTGCGGCTGCGATCCACTCCGCG 627

361 CCAAGGCGGCTACTACTTGTGACCGGCAATTTGCCAATTTGAGTCTCGGCGCATG 420
626 CCGGCGGCGCAACCTTGTCTCGGCTGCGGCTGCGGCGCTGCGATTCGCTCCCAAG 567

421 GTGGATTAATCTGAGAGTGGGCGCTGCCATGATTTGTAGCGATCTTCCAGAAATT 480
566 GTGGATTAATCTGAGAGTGGGCGCTGCCATGATTTGTAGCGATCTTCCAGAAATT 507

481 GACATTTTGGCATCACTTACGATCTGTTAAGCATCTCTATGTGTAGTGTGGGCG 540

```

/organism="Saccharum officinarum"
/mol_type="mrna"
/db_xref="taxon:4547"
/clone="SCCCCA1001.P04"
/lab_host="XlBlue MRF"
/clone_1b="C14"
note="Organ: Poot of sugarcane calli submitted to low

```

QY	402	TTTGGACTCGGTCCCAATGATGGATTAATGAGAGATGAGGCGGCGGCGGCAATGATGGTAG	461
Db	890	TATGATATTCATTCATTAATTAATGTTCCTTTGGGCAAGATGACACCTCGTTGATAGATTGA	831
QY	462	CGATGCTTCCAGAGAAATTCACATTTTGGCATCACTTATCCGATGGTTAAAGACTCCCTA	521
Db	830	CGATGCTTTTAAAGATGCGACATGGTGGGGAATTTGGCGACCGGTGGTTAACACAAATTT	772
QY	522	TGAGTACGTAGAGCGCGGATATATGCTCCATTTGTTATCGAGGCTTTCATCTGTGTAG	581
Db	771	TTCTGTTTAACCAAAACGGAACACATGCCGCGAGTGCTGTAATAAAGCTTTCTGTGGTGGGCG	712
QY	582	CACCGGTGTCGCCGGGCGGTTTGTGATCGATATTTCCACAGATGTGGGCTTGAAGAAAT	640
Db	711	AAGTGGCGCTGAGACCAAGTATGCTGTTATTCGAAAGAAATATCTTAATCCGCGGAA	652
QY	641	GTAAGTACATTTCCCTCGACCCCGGTGACGTTATATCAACGGGATATCGGCCCAAGTTA	700
Db	651	CCAAATTAACCTAAATGTTGGCGGAGATGGTACGATATGCTTCTCATTTCCACTATCTA	592
QY	701	AAGTAAATCCCGCAAAATTAATGCGCATTTGCAATTTGTGAGACAGGCCAGAAATCCCT	760
Db	591	CCGGAACATAAAGGCAAAATTAACGCTGCTTGCAAAACGCTGTATACGACAAAAAAACCGG	532
QY	761	TGCTCTACGTAGAGGAGGCGCATGCGCCCAATGCCATGCCAGGTGAGGAATTTG	820
Db	531	TTGTCTACGTAGAGCGGAGGCGGCAATCAACGCGGCGTCCCATCAGCAATTAAGAAACGG	472
QY	821	CGAAGAGTTCCAGTTGGCGGTAACAACAACCTGATGGGAATTTGGGGCTTTTGACGAA	880
Db	471	TGAGAGCGTGAAATCTGCCGTGTTTGTCTCATTAATGATGGGCGTGGGGCGCTTTCGGCAA	412
QY	881	ACCATCCCTTTGGTGGGTATGTTGGATATGATGGCCACCGCTATGCCAATTTGCCG	940
Db	411	CGCATGTCAGGCACTGGGCATGCTGGGAATGCACGCTACTACGAAGCCAAATATGACGA	352
QY	941	TCAACGAATGTAATTTGTGATTTGCAGTGGGGGCGCTTTCAGACACCGGATTAATGCGCA	1000
Db	351	TGCATTAACGCGAATGATGATTTTCGCGGTCCGGGTACATTTGATGACCGAAGCAGAAC	292
QY	1001	AACCTAGACGAATTTGCTAGCGCGGCAAAATTAATTCATTTGATCATGACCCGCGGAGG	1060
Db	291	ATCTGGAAAGTACTGCGCAATATGCACATCTTTCGATATGCAATTTGATGATCTTAATTTCCA	232
QY	1061	TGGGAAAAAACAAGGCGCTCCCGATGTGCCATTTGTGGGGGATGTAACGCATGTTTAGAAC	1120
Db	231	TTTCTATAAACCGTACGCTCGCGGATATCCGATTTGTGGGGAATGCTCGGCAAGTCTCGAAC	172
QY	1121	AGCTTTTGACAGCGGCGCCGGGAATTTGGAATTAACCCACCATATCCCATACCAATCCAGGCAT	1180
Db	171	AAATGCTTGAACCTTTGTCCCAAGAAATCCCGCATCAACCATCGAATGAGATCGCACT	112
QY	1181	GGATTAATGCATTTGATGATGGGGAACGAATTAACCCCTCCAGGTGCCCACTATGAGG	1240
Db	111	GGTGGCGCAAAATTTGAACAATGGGCGGCTGTGTAGTGCTGGAATATGACACTACACTG	52

QY 1241 ATACTATTGCCCCCAGAGGT 1262
DB 51 AAAAGATTAAACCGCAGCGCT 30

RESULT 5
BH390390/c
LOCUS
DEFINITION BH390390 826 bp DNA linear GSS 11-DEC-2001
AG-ND-153F18, TR ND-TAM Anopheles gambiae genomic clone
AG-ND-153F18, genomic survey sequence.
ACCESSION BH390390
VERSION BH390390.1 GI:17336531
KEYWORDS GSS.
SOURCE Anopheles gambiae (African malaria mosquito)
ORGANISM Anopheles gambiae
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Preygota;
Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae;
Anopheles.
1 (bases 1 to 826)
Hong,Y.S., Hogan,J.R., Wang,X., Sarkar,A., Sim,C., Loftus,B.J.,
Ren,C., Huff,B.R., Carlile,J.L., Black,K., Zhang,H.-B.,
Gardner,M.J. and Collins,F.H.
Construction of a BAC library and generation of BAC end
sequence-tagged connectors for genome sequencing of the African
malaria mosquito Anopheles gambiae
Mol. Genet. Genomics 268 (6), 720-728 (2003)
22542063
12655398
COMMENT Other GSSs: AG-ND-153F18.TF
Contact: Brendan J Loftus
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0208
Fax: 301 838 3543
Email: b.loftus@igr.org
This clone is from an A. gambiae BAC library (ND-TAM) provided by
F.H. Collins and sequenced by The Institute for Genomic Research
(TIGR). The BAC library was generated from A. gambiae PBST strain
DNA. All DNA was extracted from newly hatched first instar larvae
to minimize the inclusion of DNA from microorganisms that inhabit
the gut. The DNA is derived from mixed sexes of larvae. The BAC
library was constructed at Texas A&M University BAC Center
University College Station, Texas 77843-2123, USA using a HindIII
partial digest.
Seq primer: M13 Rev
Class: BAC ends.
Location/Qualifiers
1..826
/organism="Anopheles gambiae"
/mol_type="genomic DNA"
/strain="PBST"
/db_xref="taxon:7165"
/clone="AG-ND-153F18"
/clone_1fb="ND-TAM"
/note="Vector: pECBAC1; Site_1: HindIII"

ORIGIN
Query Match 9.0%; Score 176.8; DB 28; Length 826;
Best Local Similarity 53.0%; Pred.No.6.1e-40;
Matches 425; Conservative 0; Mismatches 372; Indels 5; Gaps 2;

QY 302 CCGCGGATGGTATGCGACGACGACGATGAAGTGGAGTTGTTGGTGGTACATCTGAC 361
DB 826 CAGCGCAGGATATGCCCCCTACGATGAAAAAGTGGGTTCGTTTGCACATACCGTTC 767

QY 362 CAGGGGCGACTACTTGTGTGACCGGCAATGGCAATGCCATTTGGACTCGGTGCCA--T 419
DB 766 CCGGTGCAACCAATTAGTAAACGATTTGGCTGATGCTCAGATGATTCAACACACACTTT 707

QY 420 GGTGGTATTACTGGAGAGGTGGGCGCGGCCCATGATTTGTGACGATGGTTTCCAGGAAT 479
DB 706 GTATGATTATCCGGGCAAGGTATTGGACATCTTTTGAACAGATGCAATTTCCAGGAAC 647

QY 480 TGACATTTTGGCATCCTTACCGATCGTTAAGCATCCTCTATGTGTAGTGGCG 519
DB 646 TGATGATCATATGTAATGATGCAAGTACCAAGTGAATTAATCATCTAATCTATC 567

QY 540 GGAATGAGCTCGCATTTGTTACTGAGGCTTTCATCTTGTAGACCGGTGTCGGGACC 599
DB 586 ACAGCTCCGACGGGTTTACCAAAAAGCATTTTATGCGCGGTTCCGGAAGACAGGACC 527

QY 600 GGTTTATGATATTTCCCAAGATGTGGCTTGAAGAAATGATGATCACTTCCCTCGA 659
DB 526 TGATCATATGATATTTCAAAAAATGCCAATTGCAACAATTTGAATA--TCCGGGCTA 470

QY 660 CCGCGTGAAGTTATCTACCGGTTATGCGCCCAAGGTTAAAGTAAATCCCGACAAAT 719
DB 469 TAGCAATGTGACATTTAGCAATTTGCGCCGGAAGCCAAATTTGAAGAAACAAATATAT 410

QY 720 TAATCGGCAATTGCATTTGTGAGACAGCGCAGAAATCCCTTGCTACGTAGGGGAGG 779
DB 409 TGAAGAGAGCTGCTGAATTTATTAATGAACCCAAAGCCATTTGTATCTTTGACAGAGG 350

QY 780 GGGATGCGCGCAATGCGCATGCCAGAGTGCGAGGAATTTGGGAAAGTTCCAGTTGCC 839
DB 349 AGTCATCTTGGAAGAACAGAAAAAGCTTTCATATTTATGAAAAGGCAATTTTCC 290

QY 840 GGTACACACCAACCTGATGGAATTTGGGAGCTTTTGAACGAAACCATCCCTTCGTGGG 839
DB 289 GGCTGCATGACATATTTAGGGGAGAGTGCATTTCTACTGATCATCGTTAGAGATTGG 230

QY 900 TATGTTGGGATATGATGCGCAACCGCTATGCGCACTTTGCCGTAGCGAATGTATTTGTT 959
DB 229 TATGTTGGGATATGATGATGATTAATGACGACCAAAATGTTTAAACAGAAATGTATGTTT 170

QY 960 GATTCAGTGGGGGCCGTTTCAGACGACCGGATGTAATGCAACTGACGAATTTGGTAG 1019
DB 169 AATTGCTATAGGAATGAGTGTTCACACCGCTGTAACGAGCAATTGACAAATATGCAAA 110

QY 1020 CCGGCGCAAGTAATTCATTCATGATGATGACCCCGCGAGGTGGAAAAACAGGGCTCC 1079
DB 109 ACAGGCGACAGTATTCATCTGATGATGATGATCTCTGCCAAGTGAATTAATGTAATAAC 50

QY 1080 CAAATGCGCATGTTGGGGGAT 1101
DB 49 TACCGTTCCGCTTTGGGGCGAT 28

RESULT 6
BG595301
LOCUS
DEFINITION BG595301 768 bp mRNA linear EST 07-MAR-2003
EST493979 cSTS Solanum tuberosum cDNA clone CTS9P16 5' sequence,
mRNA sequence.
ACCESSION BG595301
VERSION BG595301.1 GI:13613441
KEYWORDS EST.
SOURCE Solanum tuberosum (potato)
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum.
1 (bases 1 to 768)
van der Hoeven,R., Bezzerides,J., Sun,H., Cho,J., Chienlungo,A.,
Bougrit,O., Buell,C.R., Romling,C., Tanksley,S. and Baker,B.
Generations of ESTs from sprouting potato eyes
Unpublished (2000)
TITLE
JOURNAL
COMMENT Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@igr.org
This clone can be obtained from the University of Arizona Genomics
Institute. Orders can be made through URL:
http://genome.arizona.edu/orders/
Seq primer: M13P-R
Location/Qualifiers

source 1..768
 /organism="Solanum tuberosum"
 /mol_type="mRNA"
 /cultivar="Kennebec"
 /db_xref="taxon:4113"
 /clone="cSTS9P16"
 /tissue_type="sprouting eyes from tubers"
 /dev_stage="12-14 weeks post harvest"
 /lab_host="SOLR"
 /clone_lib="cSTS"
 /note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2: XhoI; Various sizes of sprouting eyes (2mm to 15mm) were taken from tubers. The tubers were incubated at 26C in the dark for 2-3 weeks prior to sprouting. The eyes were frozen in liquid nitrogen immediately upon removal from tubers."
 ORIGIN
 Query Match 9.0%; Score 176.6; DB 12; Length 768;
 Best Local Similarity 52.8%; Pred. No. 6.8e-40;
 Matches 408; Conservative 0; Mismatches 359; Indels 6; Gaps 1;
 QY 259 ATTGAGCATTTTGTGGCCGATGAAACAGAGCTTCCATGGGGGGATGGGTATGCC 318
 DB 2 ATTGCTAATGCTGCGACGTCATGAAAGGGGTGTGTGTTGCTGCAAGGGTTACGA 61
 QY 319 AGAGCCACAGGTAAAGTGGAGTTGTTCCGTACATCTGACCCAGGGCGACTACTTG 378
 DB 62 CGGGCTACTGGGTTCCCGTGGTGTTCATCTCTACATCTGCTCCGGAGCTAGAACTT 121
 QY 379 GTGACCGGCAATTGCCAATGCCATTTGGACTCGGTGCCATGTGTGATTTACTGGAG 438
 DB 122 GTTAGCGGCTTGGCGAGTCTTGTGTGATGATTCGATGCTGCTATTTACCGGTC 181
 QY 439 GTGGGCGGTGCGCATTTGTAGAGATGCTTTCAGAGAAATGACATTTTGGCATACC 498
 DB 182 GTGCCGAGAGAGATGATGTGACGATCCGTTTACAGAACTCTATTTGTAGAGTACG 241
 QY 499 TTACCGATGTTAAAGACTCTATGTGTGATGAGTGGCGGATGATGCTTCGATGTT 558
 DB 242 AGATTCATTAAGAGCATTAATTAATCTTATGATGATGAGGATATTCCTAGGGTGT 301
 QY 559 ACTGAGGCTTCCATCTTGTAGACACGGGTGTCGGGGCGGTTTGTATGATATTC 618
 DB 302 CGTGAAGCGTTTTCATGAGCAATCGGACGGCTGACCGGTTCTGATTTGATTCCT 361
 QY 619 AAGATGTGGCTTGAAGATGTGATGATTCCTCCGACCCCGGTGACGTTAATCTA 678
 DB 362 AAGATATTACAGAACTATGATCTAATGATGGATGACCAATGAGGTGCTGCT 421
 QY 679 CCGGTTATCGCCCAAGGTAAAGTAAATCCCGAATAATTAATGGGATTCGATG 738
 DB 422 TACATGTAGTGTGCTTAATGCTTAATGATGATGCTTTTGAACAAATGTTAGGCTG 481
 QY 739 TTGAGCAGGCGCAAAATCCCTGCTACGTAGGGGAGGGCGCATCGCCCAATGCC 798
 DB 482 ATTTCAAGTGAAGAGAGCCGTTTGTATGTGGGTGTGGGTGTA-----CACATCG 535
 QY 799 CATGCCAGGTGACAGAAATTTGGCGAAAGTTCCAGTTGCGGTTAACAACCCCTGATG 858
 DB 536 AGTAGAGAGTGAAGCAATTTGTGAGCTTAACGGGATTCCTGTGGGAGATCTTGTATG 595
 QY 859 GGAATTTGGGGCTTTTGAAGAAACCATCCCTTCGCTGGGTATGTTGGGTATGACATGGC 918
 DB 596 GGTCTTGGAATTTTCCATGTGGGAGATGAGCTTCTCTCAATGTTGGGTAATGATGG 655
 QY 919 CACCGCTATGCAACTTGTCCGTGACGAAATGATATTGTTGATGATGAGTGGGGCCCGT 978
 DB 656 ACTGTGATGCTATTTATGCGGTGATAGTATGATTTGTTGCTTGTGATTTGGGGTGAAG 715
 QY 979 TTGACGACCGGGTAACTGCGCAACTAGAGAAATTTGCTGACCGCGCCCAAGT 1031
 DB 716 TTGATGATGAGTTACTGTTAAATGGAACCTTTGCTAGCGGGCGAATTT 768

RESULT 7
 BG599315 775 bp mRNA linear EST 07-MAR-2003
 LOCUS EST04215 cSTS Solanum tuberosum cDNA clone cSTS25A13 5' sequence,
 DEFINITION mRNA sequence.
 ACCSSION BG599315 GI:13619156
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Solanum tuberosum (potato)
 Solanum tuberosum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 asterids; lamids; Solanales; Solanales; Solanales; Solanum.
 1 (bases 1 to 775)
 van der Hoeven,R., Bezzerides,J., Sun,H., Cho,J., Chiemingo,A.,
 Bougri,O., Bueli,C.R., Romning,C., Tanksey,S. and Baker,B.
 Generations of ESTs from sprouting potato eyes
 Unpublished (2000)
 Contact: Robin Bueli
 The Institute for Genomic Research
 9712 Medical Center Dr, Rockville, MD 20850, USA
 Email: potato-array@igr.org
 This clone can be obtained from the University of Arizona Genomics
 Institute. Orders can be made through URL:
 http://genome.arizona.edu/orders/
 Seq primer: M3F-R.
 location/Qualifiers
 1..775
 /organism="Solanum tuberosum"
 /mol_type="mRNA"
 /cultivar="Kennebec"
 /db_xref="taxon:4113"
 /clone="cSTS25A13"
 /tissue_type="sprouting eyes from tubers"
 /dev_stage="12-14 weeks post harvest"
 /lab_host="SOLR"
 /clone_lib="cSTS"
 /note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2: XhoI; Various sizes of sprouting eyes (2mm to 15mm) were taken from tubers. The tubers were incubated at 26C in the dark for 2-3 weeks prior to sprouting. The eyes were frozen in liquid nitrogen immediately upon removal from tubers."
 ORIGIN
 Query Match 9.0%; Score 175.6; DB 12; Length 775;
 Best Local Similarity 53.4%; Pred. No. 1.3e-39;
 Matches 395; Conservative 0; Mismatches 339; Indels 6; Gaps 1;
 QY 356 CTGACCAAGGGCGACTACTTGTGTGACCGGATTCGCAATGCCATTTGGACTCGGTGC 415
 DB 2 CTGGTCCGGGAGCTGCAATCTTGTAGTGTCTTGGCGAGTCTTGTGATGATATTC 61
 QY 416 CCATGTGTGATTTCTGAGAGAGTGGCGGCGTGCATGATGTAAGCATCTTCCAG 475
 DB 62 CGATTTGCTATTTACGGGTCAAGTGCAGAGAGATGATGATGATGATGATGATGATG 121
 QY 476 AATTTGACATTTTGGATCACTTACCGATGTTAAGCACTCTATGTGTGATGATG 535
 DB 122 AAACCCATTTGTGAGTAAAGATCTAATTAAGCAATTAATTAATTAATTAATTAATG 181
 QY 536 CCGCGGATATGCTTCGATTTTACTAGAGGCTTTCATCTTGTGTAAGACACCGGTGCTCCG 595
 DB 182 TAGAGGATATTCCTAGGGGTGTGTGTAAGGCTTTTCTTACGCAAAATCGGAGCGCTG 241
 QY 596 GCGCGGTTTATGATATTCCTCAAGATGTGGCTTGAAGAAATGATGATGATATTCGCC 655
 DB 242 GCGCGGTTTATGATATTCCTCAAGATGTGGCTTGAAGAAATGATGATGATGATGATG 301
 QY 656 TCGACCCCGGTGACGTTATCTACCGGGTATTCGCCCAAGGTTAAAGTAAATCCCGAC 715

Db 302 ATCAGCCATGAGTTCCTGCTGATCATGTCTAGGTACCTTAATTCCTTAATGAGATGC 361
QY 716 AAATTAATGCGGATTTGCAATTTGAGAGACAGCCAGAAATCCCTTGCTACGATGAGG 775
Db 362 TTTTGACAAATTTATGAGGCTGATTTGAGAGTGAAGAAGCCTGTTTGTATGTGGGTG 421
QY 776 GAGGGCCCATGCGCCCAATGCGCCATGCGCCAGGTGACAGAAATTTGGGAAAGTTCCAGT 835
Db 422 GTGGGTGTTTGCATCAAGTAGG-----AGCTGAGACGATTTGTGAGCTTTACGGGTA 475
QY 836 TGCCGGTATACACACCCCTGATGAGGATTTGGGGCTTTTGAGCAAAACCATCCCTTTCCG 895
Db 476 TTTCTGTGGCGAGTACTTTGATGAGGTCTTTGAGCTTTTCCAACTGGGATGAGCTTTCC 535
QY 896 TGGGATGTTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 955
Db 536 TTCAATATGTTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 595
QY 956 TGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1015
Db 596 TGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 655
QY 1016 CTAGCCGCGCCCAATTAATTCATGATGATGATGATGATGATGATGATGATGATGAT 1075
Db 656 CTAGCCGCGCCCAATTAATTCATGATGATGATGATGATGATGATGATGATGATGATGAT 715
QY 1076 CTCCCGATGTCCTCATTTGTG 1095
Db 716 ACCTCATGTTTCATTTGTG 735

RESULT 8
CC722291 811 bp DNA linear GSS 19-JUN-2003
DEFINITION OG08B62TV ZM 0.7.1.5_KB Zea mays genomic clone ZM8MA0697K04,
genomic survey sequence.
ACCESSION CC722291
VERSION CC722291.1 GI:32127067
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoidae; Andropogoneae; Zea.
Whiteaw,C.A., Quackenbush,J., Van Aken,S., Uterback,T.,
Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
Citek,R.W., Nuneberg,A., Robbins,D. and Lakey,N.
Consortium for Maize Genomics
Unpublished (2002)
Other GSSs: OG08B62TH
Contact: Cathy Whiteaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whiteaw@tigr.org
Seq primer: TP
Class: sheared ends.
Location/Qualifiers
1..811
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZM8MA0697K04"
/clone_lib="ZM 0.7.1.5_KB"
/note="Vector: pBSCS-, Site 1: HindIII, 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN
Query Match 8.9%; Score 175.2; DB 29; Length 811;
Best Local Similarity 53.5%; Pred. No. 1.8e-39;

Matches 413; Conservative 0; Mismatches 353; Indels 6; Gaps 2;
QY 275 TGCGCATGACAGAGAGCTTCCATGCGCGGATGGATATGACAGACAGGATTAAG 334
Db 45 TCGGCACAGACAGAGGAGAGGCTTTCGCGGCTCCGCGCTACAGGCGCTCTCGGGCGGG 104
QY 335 TGGAGATTGTTTGGTATCATCTGACACAGGGGCGATTAATTGTTGACCGGATTTGCCA 394
Db 105 TTGGCTCTGACATCGCACCTTCGCGCCCGCGGCCACCAACCTAGCTCTGCGCTGCGAG 164
QY 395 ATGCCCATTTGACCTGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 454
Db 165 AGCGGTGCTGACCTTCGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 224
QY 455 TTGGTACGATGCTTTTCCAGAAATTTGACATTTTGGCATCACTTACCGATGTTTAAGC 514
Db 225 TTGGCACCGACGCTTTTCCAGAGACGCCCATGCTGAGGTCAACCGCTCATACCAAGC 284
QY 515 ACTCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 574
Db 285 ACAACTACCTGCTGCTGACGATGACGACATCCCGCGCTCGTCAAGAGGCTTTTCC 344
QY 575 TTGCTAGCACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 634
Db 345 TCGCATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 400
QY 635 AAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 694
Db 401 CAGCAATGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCT 460
QY 695 CGGTTAAAGGTAATCCCGGCAATTAATGCGGATTTGCAATTTGTTGAGACAGGCGAGAA 754
Db 461 CTTCGCAAGCTCTCCCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGG 520
QY 755 ATCCCTTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 814
Db 521 CGCGCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 578
QY 815 AATTGGCGGAAAGGTTCCAGTTCGCGGTAACACACCGGATTTGAGGATTTGGGCTTTTG 874
Db 579 GCTTTGGAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 638
QY 875 ACAGAAACCATCCCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 934
Db 639 CCAAGCAGACCCACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 698
QY 935 TTGCGCTGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 994
Db 699 ATCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 758
QY 995 CTGCAAACTAGACGATTTGCTAGCGCGCCCAATTAATTCATTTGACAT 1046
Db 759 CAGGAAATTTGAGGCTTTTTCAGGACAGAGCTTAAGTTTGACATTTGAT 810

RESULT 9
CK166070 1121 bp mRNA linear EST 05-DEC-2003
DEFINITION FGAS050122 Triticum aestivum FGAS: TaLc7 Triticum aestivum cDNA,
mRNA sequence.
ACCESSION CK166070
VERSION CK166070.1 GI:38998757
KEYWORDS EST.
SOURCE Triticum aestivum (bread wheat)
ORGANISM Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Triticum.
1 (bases 1 to 1121)
Allard,F., Crosby,W.L., Danyluk,J., Endes,F., Frick,M., Gaudet,D.,
Genswein,B., Graf,R., Gulick,P., Hyman,L.D., Laroche,A.,
Links,M.G., McCarthy,E.L., Monroy,A., Muzak,I., Nilsson,D.,
Penniket,C., Roach,J.L. and Sarhan,F.

TITLE Functional Genomics of Abiotic Stress in Wheat and Canola Crops
JOURNAL Unpublished (2003)
CONTACT: Mm L Crosby
COMMENT Bioinformatics

University of Saskatchewan, Department of Computer Science
1101 Engineering Building, 57 Campus Drive, Saskatoon,
Saskatchewan, S7N 5A9, Canada
Tel: 306 966 1769
Fax: 306 966 2033

Email: fgsa@scs.usask.ca
This sequence is the direct result of the Base calling software
Phred (default parameters). It is the raw base calls. No aid in the
identification of the high quality insert the software Lucy
(default parameters) has been run on this sequence. Lucy identified
the region [63,872].
Plate: Talc707 row: J column: 19.

FEATURES

source

1. 1121

/organism="Triticum aestivum"

/mol_type="mRNA"

/cultivar="Wheat line CI 14106"

/db_xref="taxon:4565"

/lab_host="DH5 alpha"

/clone_id="Triticum aestivum FGAS: Talc7"

/note="Organ: Crown; Vector: pGEM-T; SSH (suppression
subtractive hybridization) cDNA library from genotype
CI14106 cold hardened at 2 C for 21 days and 49 days
(equal amount of cDNA pooled together before subtraction,
tester) and subtracted against genotype CI14106
non-hardened (20 C) (driver). Nitro-pyrole anchored
oligo-dT priming and non-directional cloning."

ORIGIN

Query Match

Best Local Similarity 8.9%; Score 174.6; DB 14; Length 1121;
Matches 423; Conservative 0; Mismatches 349; Indels 9; Gaps 3;

QY 806 AGGTCAGAGAAATTCGCGAAAGTTCCAGTCCGCGAAACAACACCTGATGGAAATTG 865
DB 212 AGTTGCGCGCGCTTGTGAGCTTACCTGGAATCCAGTTACATCACTCTATGGCCTTG 271
QY 866 GGGCTTTGACGAAAACCATCCCTTTCGGGGTATGTTGGTATGATGATGATGATGATG 925
DB 272 GCAACTTCCCGACGACGACCATCTGCTGCGCATGCTGGGATGATGATGATGATGATG 331
QY 926 ATGCAACTTTCGCGTACGAGAAATGATTTGTTGATGATGATGATGATGATGATGATGATG 985
DB 332 ATGCAAAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 391
QY 986 ACCGGGTAAGTGGCAAACTAGACGAATTTGCTAGCCGCGCAAGTAATTCACATTGACA 1045
DB 392 ATCGTGTGACTGGGAAATCGAGGCTTTGCAAGCAGATGCAAGATTGTCACATTGACA 451
QY 1046 TCGACCCGCGGAGGTGGGAAAAAACAAGGCTCCCGATGCGCCATTGTTGGGGATGTAC 1105
DB 452 TTGACCCAGCTGAAATGGCAAGAACAGACCATGCTCTCATTTGTCAGATGTGA 511
QY 1106 GCCATGTTTTGAACACGCTTTTTCAGCGGGCCCGGAAATGATACCCCAACCATCCCC 1165
DB 512 AGCTTGCTT---ACAGGGGTGAATGATCTATTAATGGAGCAACACAGAGGTC 568
QY 1166 ATACACCCAGGAGTGTGAATGATGATGATGATGATGATGATGATGATGATGATGATG 1225
DB 569 TGAATTTTGTTCATGACACAAGAGTGTGATGACAGAAAGAGGATTTCTCTTGAGAT 628
QY 1226 TGCCCACTATGAGATATGATTTGCCCCAGAGGAGTATGACAGAAATGATGATGATGATG 1285
DB 629 TCAAGACTTTTGGGAGGCAATCCGCGCAATATGATATGATGATGATGATGATGATGATG 688
QY 1286 C---CCCCGATGCTACTATGACACGATGTTGGAGCAACCAATATGTTGGGCGCCAGT 1342
DB 689 CAAAGGGAGGCGCATGTCCTGCTGTGTGGGCGAGCACGATGTTGGGCGGCTCAGT 748

QY 1343 TTTTGA---ACAATGCCCCCGCCGATGATTTTCAGTGTGCTTGGGTATGATGGCT 1399
DB 749 ATTACACTTACAGAGGGGACAGGCGAGTGCCTGCTTCTGCTGTTGGGGCAATGGAT 808
QY 1400 TTGTTTACTTCCCGCATGGAGCAAGTGGAGTGGGAGCAGCGCTCATTTGCAT 1459
DB 809 TTGGGTTACCACTCAGCTGAGCGCTGCTGCGCCACCAAGTGTACAGTTGTTGAT 868
QY 1460 CAGTGAAGTGCACCTTCCAAATGATCTTCAGAGACTGGGAAACCTAGCCAGTACGA 1519
DB 869 TGATGTATGATGATTTTCCCTCATGACATTCAGAGATGGCGTTGATCCGCAATTGACA 928
QY 1520 CATTCAGGTTAAACTATTTATTCATATGAGGTTGGCAGGGGATGTGCTCATGTGCA 1579
DB 929 CCTCCAGTTAGGTTGATATATTGAACACACAGCATTTGGGAAATGTGGGCTCGGGGA 988
QY 1580 A 1580
DB 989 A 989

RESULT 10

CB626704

OS11EB01C05.f OS11EB Oryza sativa (indica cultivar-group) cDNA
DEFINITION clone OS11EB01C05.5', mRNA sequence.

ACCESSION CB626704.1 GI:29621693
VERSION
KEYWORDS
SOURCE
ORGANISM

Oryza sativa (indica cultivar-group)
Oryza sativa (indica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE

AUTHORS

1 (bases 1 to 812)
Jantanasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,
Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.
large-scale identification of ESTs involved in the interaction
between rice and Magnaporthe grisea

JOURNAL

COMMENT

Unpublished (2003)
Contact: Rod Wing
Arizona Genomics Institute
University of Arizona
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
85721-0088, USA
Tel: 520 626 3967
Fax: 520 621 9288
Email: http://genome.arizona.edu

PCR Primers
FORWARD: gta aac cga cgg coa gtc
BACKWARD: gga aac agc tat gac cat g
Plate: 01 row: C column: 05
Seq primer: gta aac cga cgg coa gtc.

FEATURES

source

1. 812

/organism="Oryza sativa (indica cultivar-group)"

/mol_type="mRNA"

/cultivar="IR36"

/db_xref="taxon:39946"

/clone_id="OS11EB01C05"

/issue_type="leaf"

/dev_stage="3 week"

/lab_host="DH10B"

/clone_id="OS11EB"

/note="Vector: pBluescript II KS +; Site 1: EcoRI; Site 2:
XhoI; 24 hrs after inoculation with Rice Blast (PO6-6-3)"

ORIGIN

Query Match 8.9%; Score 173.2; DB 14; Length 812;
Best Local Similarity 54.0%; Pred. No. 6.9e-39;

Matches 380; Conservative 0; Mismatches 318; Indels 6; Gaps 1;

QY 415 CCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 474

```

Db      1 CCGATGGTCGCATACCGGCGCAGGTCCCGCGCGCATGTCGGACCGAGCCCTTCAG 60
Qy      475 GAATTTGACATTTTGGGACATCTTACCGATCGTTAGCACTCTATGTGTAGT 534
Db      61 GAGAGCCCATAGTGTAGTGTACCCCGCTCCATCACAAGACATTAATCTGTCTTAT 120
Qy      535 GCGCGCATATGCTGCAATTTGTACTAGAGCTTTCATCTTCTAGCAACGCTGCTCC 594
Db      121 GTGGAGGACATCCCGCGGTATACAGAGGCTTCTCTCGGCTCTCGGCGCTCT 180
Qy      595 GGGCGGTTTGTATGATATTCACCAAGATGTGGCTTAGAAGATGTGATTAATCC 654
Db      181 GCGCGGTGCTGTGTGACATCCCAAGACATCCAGCAGATGGCTGTCCAGTCTGG 240
Qy      655 CTCGACCCCGGTGACGTTATCTACCGGTTATGCGCCCGACGTTAAAGTATCCCGA 714
Db      241 GACACCTCGATGAATCTACCGGGGTACATTGACAGCTCGCCCAAGCCACCGCGACA 300
Qy      715 CAATTAATGGGCGCATTTGCAATTTGAGAGAGGCAAAATCCCTTCTTACGTAGG 774
Db      301 TTGCTTGAAGAGGCTTCTGCTGTGTGGAGTACGCGCCCGATTTCTATGTGCT 360
Qy      775 GAGAGGCGCATGCGCCCAATGCCCATGCGAGTGACAGAAATTTGCGAAAGTTCAG 834
Db      361 GGTGG-----CTGCTTGCACTGTGTATGATTAATGCGCGCTTTGTGAGCTACCGG 414
Qy      835 TTGCGGTAAACAACACCTCTGATGGGAATGGGCTTTTGAAGAAACATCCCTTGG 894
Db      415 ATCCCAATTAACAACACTGTATGGGCTCGGCAATTTCCAGTATGATCGTGTGCC 474
Qy      895 GTGGGTATGTGGGTATGATGATGAGCCACCGCATATGCACTTTGCGGACAGAAATGAT 954
Db      475 CTGCGCATGCTTGGATGATGATGACAGGCTGTACGAAATTAATGCGTGTGATAGGCTGAC 534
Qy      955 TTGTTGATTCAGATGGGGGCGCGTTTGCAGCAACCGGTTACTGTGCAATAGCAATTT 1014
Db      535 CTGTTGCTTGCATTTGGGCGGTGTTGATGATCGTGTGACAGGAAATTTAGGCTTTT 594
Qy      1015 GCTACCCCGCCCAAGTATTCATTTGACATCGACCCGCGGAGGTGGAAAAACAG 1074
Db      595 GCAACAGAGGCCCAATTTGTCATTAATGATTCAGCGGGAATTTGAAAGAACAG 654
Qy      1075 GCTCCGATGTCGCCATTGTGGGGATGTAGCGCATGTTTGA 1118
Db      655 CAACCAATGTGTCAATTTGGCGCATGTTAAGCTTGCTTACA 698

```

RESULT 11
 CG970807 897 bp DNA linear GSS 15-DEC-2003
 LOCUS MBENF13TF mhz Medicago truncatula genomic clone 7C2, genomic
 DEFINITION survey sequence.

ACCESSION CG970807.1 GI:39896586
 VERSION CG970807.1
 KEYWORDS
 SOURCE Medicago truncatula (barrel medic)
 ORGANISM Medicago truncatula
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifoliaceae;
 Medicago.
 1 (bases 1 to 897)
 TOWN, C.D., Shetty, J., Koo, H. and Feldblum, T.F.
 TITLE Sequencing of BAC ends from Medicago truncatula
 JOURNAL Unpublished (2003)
 COMMENT Other GSSes: MBENF13TR
 Contact: Chris Town

TIGR
 9712 Medical Center Drive, Rockville, MD 20850, USA.
 Tel: 301-838-3523
 Fax: 301-838-0208
 Email: cdtown@tigr.org

Seq primer: TGTAAAAACGAGCCAGT
 Class: BAC ends.
 Location/Qualifiers
 1..897
 /organism="Medicago truncatula"
 /mol_type="genomic DNA"
 /cullivar="genotype A17"
 /db_xref="taxon:3880"
 /clone_lib="7C2"
 /clone_1ib="mhz2"
 /notes="Vector: pBelobac11; Site 1: HindIII; Site 2:
 HindIII; Cook, D.R. and Kim, D.J, unpublished"

ORIGIN
 Query Match 8.7%; Score 170.8; DB 29; Length 897;
 Best Local Similarity 51.4%; Pred. No. 3.7e-38;
 Matches 456; Conservative 0; Mismatches 417; Indels 15; Gaps 2;

```

Qy      117 TGTTCAGGCAAGGCAACCGGGGCTTATATCTGATGGAATAGCTGAACGCCATGG 176
Db      873 TGGCCCTAACGAACCTCGAAAGGCGCGACATCTCTGTAATCACTGAACTGAG 814
Qy      177 GGTCAACACATTTTGGCTATCCCGCGGGGCAATTTGCCATCTATGATGAATGTA 236
Db      813 AGTTACAAATGCTTCTGCTTACCCCGCGGCGCTCATAGAGATTCAATCAAGCACTCAG 754
Qy      237 CCGCTTTGAAGCGCGGGGAAATTGACATATTTGTGTGCGCATGAACAGAGCTTC 296
Db      753 GCGCTCCACCGCC-----ATTGTAACGTCCTCCCTGTCACGAACAGGTGAAT 703
Qy      297 CCATGCGCGGATGGGTATGCCAGACCAAGGTAAAGTGGAGTTTTCGTTACATC 356
Db      702 CTTCGCTGCCGAGGATAGCTGCTGCTCTGCTGCTACCTGATGCTGATGCCCATTC 643
Qy      357 TGAACAAGGGGCACTAATCTGTTGACCGGATTCGCAATGCCATTTGACTCGGTTGC 416
Db      642 CCGCCCTGGGCGCACTAATCTTGTCAGTGACTGTGATGCTATGCTTGAACAGTGTTC 583
Qy      417 CATGTGTGATTAAGTGAAGGTGGCGCGTGCATGATGTGAGAGTGTTCACAGA 476
Db      582 GCTTATTTGCTATACCGGTGAGTTCCCGAGATGATTTGTAAGTGTGATCTTTCAAG 523
Qy      477 AATTGACATTTTGGCATCACTTACCGATCGTTAACAATCTCTATGTGTGATGTGC 536
Db      522 AACTCGATTTGTAAGTAACTGATCAATCAACAAAGCAATTAATCTTGTGATG 463
Qy      537 GCGGATATGCTGCGATTTGTAAGGCTTTCCATCTTGTAGACACGCTGTCCCG 596
Db      462 TGAATATATCAAGATATGTAAGAGGTTTATTAGCTTCTAGTGTAGACCTGG 403
Qy      597 GCGGTTTGTATGCAATTCACCAAGATGTGGGCTTAGAAGATGTGAGTATCCCT 656
Db      402 AACTGTTTGTATGATATACCTAAGATATTTGACAAAGGTTTCCCTCCGAATTGGGA 343
Qy      657 CGAACCCCGGTGACGTTATATACCGGTTATCCGCCACGCTTAAAGTATCCCGACA 716
Db      342 TCAACCTATTAAGTTAATCGGGTATATGAATAGTTTACCAGAGCCCGATAGGACA 283
Qy      717 AATTAAAGCGGCAATTCATTTGAGAGAGGCGCAAAATCCCTTCTCTACGTAAGGGG 776
Db      282 TTGGAACAGATTTGATGTTGTTATTTAGAGTCTAAGAAACCTGTTTGTATGTGGTG 223
Qy      777 AGGAGCGATGCGCGCAATGCCATGCCAGGTGACAGAAATTTGCGAAAGTTCAGTT 836
Db      222 TGGTAGTTT-----GAATGTATGAGGATTAAGAGCGCTTTGTGATTAATCGTGT 169
Qy      837 GCGGTAACAGACACCTGATAGGAAATGGGAGCTTTTGAAGAAACATCCCTTCGGT 896
Db      168 TCTGTGTCTAGTACTTGAATGGGTTGGTTTATATCTAGTTGATGAAGAAATTCGTT 109
Qy      897 GGGTATGTTGGATATGATGACCGACCGATATGCAACTTTGCGGTGAGGAATGTATTT 956
Db      108 GCAGATGCTTGGATGATGAGGACAGTTATGCTAATTAAGCGGTGATTAAGAGTATCT 49

```

QY 957 GTTGAATTCAGTGGGGGCGCCGTTTCAGACGACCGGTAATGCAAACT 1004
 DB 48 TTTCGTTGGGTTAGGTTTGAATCGGGTTACTGCGAAGCT 1
 RESULT 12
 CG851756 708 bp DNA linear GSS 18-NOV-2003
 LOCUS ZMMBB0340L06.r ZMMBB Zea mays subsp. mays genomic clone
 DEFINITION ZMMBB0340L06 3', genomic survey sequence.
 ACCESSION CG851756
 VERSION CG851756.1 GI:38378617
 KEYWORDS GSS.
 SOURCE Zea mays subsp. mays (maize)
 ORGANISM Zea mays subsp. mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoidae; Andropogoneae; Zea.
 REFERENCE 1 (bases 1 to 708)
 AUTHORS Yu, Y., Kim, H.R., Hatfield, J., Soderlund, C., Bharti, A.K., Messing, J.,
 and Wang, R.
 TITLE Sequencing of the maize genome
 JOURNAL Unpublished (2003)
 COMMENT Arizona Genomics Institute
 University of Arizona
 Biological Sciences West, 440A, P.O. Box 210088, Tucson, AZ
 85721-0088, USA
 Tel: 520 626 3967
 Fax: 520 621 9288
 Email: http://genome.arizona.edu
 PCR primers
 FORWARD: T7
 BACKWARD: M13r
 Plate: 0340 row: L column: 06
 Seq primer: M13r
 Class: BAC ends.
 FEATURES
 source
 1..708
 location/Qualifiers
 /organism="Zea mays subsp. mays"
 /mol_type="genomic DNA"
 /cultivar="B73"
 /sub_species="mays"
 /db_xref="taxon:4578"
 /clone="ZMMBB0340L06"
 /lab_host="DH10B"
 /clone_lib="ZMMBBb"
 /note="Vector: pRelBAC11, Site_1: HindIII, Site_2:
 HindIII; Zea mays L. ssp. mays"
 ORIGIN
 Query Match 8.7%; Score 170.6; DB 29; Length 708;
 Best Local Similarity 53.6%; Pred. No. 3,7e-38;
 Matches 381; Conservative 0; Mismatches 324; Indels 6; Gaps 1;
 QY 356 CTGACACGAGGGGCACTAACTTGTGACCGGCAATGCCAATTTGAGTCCGATC 415
 DB 705 CGGCCCCCGGCGCACCAACCTTGTCTCGGCTCGCGGACGCGCTCTCGATTCGCTCC 646
 QY 416 CCATGGTGTGATTAAGTGAAGGTGGCGGTGCGCATGTGATGAAGGAGCTTTCCAGG 475
 DB 645 CCATGGTGTGATTAAGTGAAGGTGGCGGTGCGCATGTGATGAAGGAGCTTTCCAGG 586
 QY 476 AAATGACATTTTGGATCACCCTTACCGATCGTTAAGCACTCTGATGTGACGATG 535
 DB 585 AGAGCCCATCGTGAAGGTACCGGCTCATCAACAGCAACTACCTGTCCTCGAAG 526
 QY 536 CGGCGGATATGGCTCGATTTGTATGAGGCTTTCCATCTTGTGACACCGGTCGTC 585
 DB 525 TCGACGATCCCGCGGTGTGACGAGGCTTTCTTCCTCGCTCTGTCGACCGG 466
 QY 596 GCGCGTTTGTGATGATTCACCAAGATGTGGCTTAGAAGATGTGATGATTCATCCCC 655

DB 465 GCGCGGTGCTTTCACATCCCAAGACATCCAGACACACAAATGCGGTGCTGTG 406
 QY 656 TCGACCCCGGTACGCTTAATACCGGGTTATCGCCACAGGTTAAAGATATCCCGAC 715
 DB 405 ACAAGCCATGAGTCTGCTGGGTACATTCGCGCTTCCAGCCCCCTCGCATGAGT 346
 QY 716 AAATTAATGCGCATTTGATTTGAGACAGGCGCAAAATCCCTGCTACGTAGGG 775
 DB 345 TCGTGAAGAGGTGCTGCGTCTTGTGATGATCCCGCGCTCTTCTTATGTGCG 286
 QY 776 GAGGGGCGCATGCCCGCCATGCGCATGCGCCAGTGCAGAAATTGCGGAAAGTTCCAGT 835
 DB 285 GTGG-----CTGCCAGCATGTGAGAGATTGCGACGCTTGTGAGCTGACTGAA 232
 QY 836 TCCCGGTAAACAACACCTGTATGAGAAATTGGGCGCTTTTGAAGAAACCATCCCTTGG 895
 DB 231 TCCCGGTAAACAACCTGTATGAGAAATTGGGCGCTTGGAACTTCCAGAGACAGACCATGCTTC 172
 QY 896 TGGGTATGTGGGTATGATGAGCCACCGCTATGCCAATTTGGCTGCGCAATGTATT 955
 DB 171 TCGCATCTAGTATGATGATGACAGGTGTATGCAAAATTATGCAATGATGAGCCGATC 112
 QY 956 TGTGATTCAGTGGGGGCGCTTTCAGACGACCGGTTACTGCAAACTAGACGAATTTG 1015
 DB 111 TGTTCCTTGCACTTGTGTGCGGTTTGTATGATCGTGTACAGGAAAGTTAGGCTTTG 52
 QY 1016 CTAGCGGCGCAAAATTAATTCATTTGACATGACATGACCGCGGAGGTGGAA 1066
 DB 51 CAAGCAGGGCTMAAAGGTGACGTTGATTTGATCCGCTGAGATTTGGCA 1
 RESULT 13
 CDS06TBR 1096 bp DNA linear GSS 05-JUN-2001
 LOCUS T3 end of clone AM0A030D06 of library AM0A from strain CLB 89 of
 DEFINITION Yarrowia lipolytica, genomic survey sequence.
 ACCESSION AL414349
 VERSION AL414349.1 GI:12187440
 KEYWORDS GSS.
 SOURCE Yarrowia lipolytica
 ORGANISM Yarrowia lipolytica
 Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 Saccharomycetales; Dipodascaceae; Yarrowia.
 REFERENCE 1 (bases 1 to 1096)
 AUTHORS Soulier, J.L., Aigle, M., Artiguenave, F., Blandin, G.,
 Bolotin-Fukuhara, M., Bon, E., Broctier, P., Casaregola, S.,
 de-Montigny, V., Dujon, B., Durren, P., Lepingle, A., Llorente, B.,
 Malpertuy, A., Neuvéglise, C., Ozier-Lafogère, O., Potier, S.,
 Saurin, W., Tekala, F., Toffano-Nioche, C., Weslowski, J.,
 Wincker, P. and Weissenbach, J.
 TITLE Genomic exploration of the hemiascomycetous yeasts: 1. A set of
 Yeast species for molecular evolution studies
 JOURNAL PERS Lett. 487 (1), 3-12 (2000)
 MEDLINE 20584711
 PUBMED 11152876
 REFERENCE 2 (bases 1 to 1096)
 AUTHORS Casaregola, S., Neuvéglise, C., Lepingle, A., Bon, E., Feyrerol, C.,
 Artiguenave, F., Wincker, P. and Galliard, C.
 TITLE Genomic exploration of the hemiascomycetous yeasts: 17. Yarrowia
 lipolytica
 JOURNAL PERS Lett. 487 (1), 95-100 (2000)
 MEDLINE 20584727
 PUBMED 11152892
 REFERENCE 3 (bases 1 to 1096)
 AUTHORS Genoscope.
 TITLE Direct Submission
 JOURNAL Submitted (07-SEP-2000) Genoscope - Centre National de Séquençage,
 2 rue Gaston Crémieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail :
 segre@genoscope.cns.fr Web : www.genoscope.cns.fr)
 COMMENT This GSS is part of a random genomic sequencing program of thirteen
 yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces
 exiguus, Saccharomyces servazzii, Zygosaccharomyces rouxi.

Db 190 CGCCCATCGTGGAGGTACCCGCTCATACCAAGACAACTACCTG9TCTCGACGTGC 249

Qy 539 CGGATATGGCTCGCATTTTACTAGAGGTTTCATCTTGCTAGACACCGGTGCTCCGGGC 558

Db 250 ACGATATCCCGCGCTCGTGCAGAGGCGTTTCTTCGATACCTCTGTCGCGCCGGGAC 309

Qy 539 CGGTTTATGATATTCCTCAAGATGCGGCTTAGAAGATGATGATACATTTCCCTCG 658

Db 310 CGGTCTTGTGTGATACCTCCCAAGACAT----CCAGACAGCATGGCGGTGCGCGCTTG 365

Qy 659 ACCCGGTGACGTATATTCACCGGTTATCGCCCAAGGTTAAAGTAAATCCCGACAA 718

Db 366 GACACGCCCATGATGCTGCTGGGTGATCGCGGCTTCCCAAGCTCCCGCATGAA 425

Qy 719 TTAATGCGGATGTCATTTGTTGAGACGACGCAAAATCCCTGCTACGTAGG3GAG 778

Db 426 TTTCTGACAGAGTGCTGCTGCTTGTGTGATACAGGCGCCCTGTTCTTATGTTGC 485

Qy 779 GGGCGATGCGCGCAATGCCATGCCCAAGTGCAGAAATTTGCGAAAGTTCCAGTTGC 838

Db 486 GTGTGCTGTGCAGCA--TCAGTGAAGGAGTTGTGCCCTTTGTGAGTTGACTGGAATCC 543

Qy 839 CGGTAAACACACCCCTGATGGGAATGGGGCTTTTGAAGAAACATCCCTTCCGTTG 898

Db 544 CAGTACACTACTCTTTATGGGCTTTGCAACCTTCCAGGACGACCTGCTGACTGC 603

Qy 899 GTATGTTGGTATGATGCGCACCGCTATGCAACTTTGCGCTGACGGAATGATTTGT 958

Db 604 GCATGCTTGTATGATGATGACAGATGATGCAATTTATGAGTGAATAGAGCCGATCTGT 663

Qy 959 TGATTCAGTGGGGGCGCTTTGACAGACCGGGTAACTGGCAAACTAGAGAAATTTGCTA 1018

Db 664 TGCTTCATTTTGTGTGGGCTTTGATGATGCTGGAACAGGAAATTTGAGGCTTTTGAC 723

Qy 1019 GCCCGCCCAAGTAACTCATTCATTCGACCCCGCGG 1057

Db 724 GCAGAGCTAATGATGACATGATGATGATCTGCTG 762

RESULT 15

LOCUS B1935831 802 bp mRNA linear EST 18-OCT-2001

DEFINITION EST55720 tomato flower, anthesis Lycopersicon esculentum cDNA

ACCESSION B1935831

VERSION B1935831.1 GI:16250303

KEYWORDS EST.

SOURCE Lycopersicon esculentum (tomato)

ORGANISM Lycopersicon esculentum

REFERENCE van der Hoeven,R.S., Bezzerides,J.L., Karamycheva,S.A., Tsai,J., Utechtack,T., Van Aken,S., Roming,C.M., Nierman,W., Fraser,C.M., Martin,G.B., Giovannoni,J.J. and Tanksley,S.D. Generation of ESTs from tomato flower tissue, anthesis (2001)

AUTHORS Unpublished (2001)

TITLE Contact: CUGI

JOURNAL Clemson University Genomics Institute

COMMENT 100 Jordan Hall, Clemson, SC 29634, USA

100 Email: <http://www.genome.clemson.edu/orders/index.html>

This clone is available through the Clemson University Genomics Institute

Seq primer: T3.

FEATURES

Source

Location/Qualifiers

1..802

organism="Lycopersicon esculentum"

molecule="mRNA"

clone="TA496"

db_xref="taxon:4081"

clone="CTOD24K17"

/tissue_type="flower"

/dev_stage="anthesis"

/clone_lib="tomato flower, anthesis"

/note="vector: pBluescript SK(-); Site_1: EcoRI; Site_2: XhoI; supplier: Cornell University; sequencing: The Institute for Genomic Research; flower buds and flowers were taken from greenhouse plants (4-8 wks old, TA496). They were immediately frozen in liquid nitrogen and then size-separated while remaining frozen."

ORIGIN

Query Match 8.4%; Score 164.6; DB 12; Length 802;

Best Local Similarity 52.4%; Pred. No. 2.3e-36;

Matches 389; Conservative 0; Mismatches 349; Indels 5; Gaps 1;

Qy 259 ATTGACATATTTTGTGGCGCCATGAACAAGAGCTTCCATGGCGCGGATGGGTATGCC 318

Db 65 ATTGATATGCTGCTACCAAGTATGAGCAAGGTGTGTTTGTGCTGACAGAGGTTACGCA 124

Qy 319 AGAGCCACAGGTAAAGTGGAGTTTGTTCGATACATCGACACAGAGGCGCACTAATG 378

Db 125 CGGCTACTGGGTTCCTGGTGTGTTGATGCTACTGCTGCTGCTGCTGCTGCTGCTGCTG 184

Qy 379 GTGACCGGATTTGCCAATGCCATTTGAGTCTGCGTCCATGCTGCTGCTGCTGCTGCTG 438

Db 185 GTTATGCTGCTGCGAATGCTTGTGATGATGATGATGATGATGATGATGATGATGATG 244

Qy 439 GTGGCCCGGCAATATTTGATGAGCATCTTCCAGAAATTTGACATTTTGGCATCAC 498

Db 245 GTCCCAAGAGATATTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 304

Qy 499 TTACCGATGTTAAGCACTCCTATGTGATGATGATGATGATGATGATGATGATGATG 558

Db 305 AGATCTATACCAACATATATTTTGTATGATGATGATGATGATGATGATGATGATGAT 364

Qy 559 ACTGAGCTTTCATCTGCTAGACACCGGTGCTCCCGGCGGTTTGAATGATATCCC 618

Db 365 CGTGAGCAATTTTCTTCCGAAATCGGACGCGCTGCGCCAGTTTGTATGATGATGAT 424

Qy 619 AAGATGTGGGCTTAGAAGAAATGATGATGATGATGATGATGATGATGATGATGATG 678

Db 425 AAGATATTTACAGCAATATGATGATGATGATGATGATGATGATGATGATGATGATG 484

Qy 679 CCGGCTATGCGCCCGCAAGGTTAAAGTATCCCGCAAAATTAATGGGCGATGCAATG 738

Db 485 TACATGCTAGATTAACCTAAATTTGCTATGATGATGATGATGATGATGATGATGATG 544

Qy 739 TTGAGCAGGCGCAAAATCCCTGCTACTAGAGGAGGAGGCGATGCGCCCATGCC 798

Db 545 ATTTCGAGTGAAGAAACCTGTTTGTATGTGGGTGTGGGTGTGGCATCA---- 600

Qy 799 CATGCCAGGTGACAGAAATTTGCGAAAGTTCCAGTTGCGGTTAACAACCATCTGATG 858

Db 601 -GTGAGGAGCTGAGACGATTTGTGAGGTTACAGATATTCCTGTAGGAGATCTTATG 659

Qy 859 GGAATTTGGGCTTTTGAAGAAACATCCCTTCCGAGGATGATGATGATGATGATGATG 918

Db 660 GGTCTTGAAGCTTTTCCACTGCGGAGTACGTTTCACTTCAAGTGTGGGTATGATG 719

Qy 919 CACGCTATGCAACTTTGCGGTCGACGGAATGATGATGATGATGATGATGATGATGATG 978

Db 720 ACTGTATGCTAATTTATGCTGTGATGATGATGATGATGATGATGATGATGATGATG 779

Qy 979 TTGACGACCGGGTAACTAGCGCA 1001

Db 780 TTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 802

Search completed: July 25, 2004, 07:28:08

Job time : 5595.05 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 25, 2004, 03:57:37 ; Search time 237.282 Seconds

(without alignments)
10115.521 Million cell updates/sec

Title: US-09-893-033-17

Perfect score: 565

Sequence: 1 gtggaatttaccacaaatg.....ccctggaatccaagtcttag 565

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124039041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_29Jan04:*
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002s:*
7: Geneseqn2003as:*
8: Geneseqn2003bs:*
9: Geneseqn2003cs:*
10: Geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	565	100.0	565	6	AAD27618
2	141.2	25.0	1721	6	ABa98820 Soybean A
3	138.4	24.5	110000	4	AA199682_33
4	138.4	24.5	110000	4	AA199683_33
5	137	24.2	504	7	ACA38392
6	137	24.2	507	7	ACA38392 Prokaryot
7	135.8	24.0	504	7	ACA40768
8	135.4	24.0	504	7	ACA37689 Prokaryot
9	134.6	23.8	1813	6	ABa98816 Corn ALS
10	133.8	23.7	492	7	ACA26110 Prokaryot
11	133.8	23.7	13732	3	AAA81484
12	133.8	23.7	110000	3	AAA81489_1
13	133.8	23.7	110000	3	AAA81489_2
14	133.8	23.7	110000	3	AAA81489_2
15	133.8	23.5	1297	6	ABa98814
16	132.2	23.4	489	7	ABa98811 N. gonorr
17	132.2	23.4	489	7	ACA41380 Prokaryot
18	129	22.6	489	7	ABa98819
19	127.6	22.6	489	7	ACA27234
20	126.6	22.4	489	7	ACA25535 Prokaryot
21	126	22.3	489	7	ACA24083 Prokaryot
22	117.2	20.4	489	6	ACA39862 Prokaryot
23	115.4	20.4	489	6	ABa98823 M. capsul

24	114	20.2	349980	6	ABa981844	ABa981844 Bifidobac
25	112	19.8	960	6	ABa98543	ABa98543 Arabidops
26	112	19.8	1673	3	AAA27424	AAA27424 Arabidops
27	111.8	19.8	492	4	AAa54309	AAa54309 Pseudomon
28	111.8	19.8	492	7	ACA42653	ACA42653 Prokaryot
29	110	19.5	1861	2	AAV55857	AAV55857 Plant ace
30	109.4	19.4	516	5	AAH68423	AAH68423 C glutam
31	109.4	19.4	639	4	AAH71477	AAH71477 Coryneb
32	109.4	19.4	2427	4	AAH54905	AAH54905 Nucleotid
33	109.4	19.4	349980	5	AAH68528	AAH68528 C glutam
34	105.8	18.7	519	7	ACA35460	ACA35460 Prokaryot
35	105.8	18.7	519	7	ACA19363	ACA19363 Prokaryot
36	104.6	18.5	507	8	ADA32571	ADA32571 DNA encod
37	104.4	18.5	522	7	ACA29659	ACA29659 Prokaryot
38	104.2	18.4	489	4	AAa51405	AAa51405 Klebsiell
39	104.2	18.4	522	4	AAa54008	AAa54008 Klebsiell
40	103.8	18.4	489	7	ACA43425	ACA43425 Prokaryot
41	103	18.2	489	7	ACA21164	ACA21164 Prokaryot
42	101.8	18.0	489	7	ACA33063	ACA33063 Prokaryot
43	100.6	17.8	495	7	ACA33434	ACA33434 Prokaryot
44	100.2	17.7	492	4	AAa53522	AAa53522 Haemophil
45	100.2	17.7	492	7	ACA34494	ACA34494 Prokaryot

ALIGNMENTS

RESULT 1	
AAAD27618	
ID	AAAD27618 standard; DNA; 565 BP.
XX	
AC	AAAD27618;
XX	
DT	18-APR-2002 (first entry)
XX	
DE	Synechocystis sp. AHA5 small subunit gene, 111N.
XX	
KM	Cyanobacteria; acetylhydroxyacid synthase; AHA5; acetylacetyl synthase;
KM	phytoene desaturase; PDS; plant trait control; plasmome transformation;
KM	herbicide resistance; fungicide resistance; insecticide resistance;
KM	plant growth protectant; ds.
XX	
OS	Synechocystis sp.
XX	
FN	W0200200915-A2.
XX	
PD	03-JAN-2002.
XX	
PF	27-JUN-2001; 2001MO-US020338.
XX	
PR	27-JUN-2000; 2000US-0214705P.
XX	
PA	(AMCV) AMERICAN CYANAMID CO.
XX	
FI	Kakefuda G, Koop H, Sturmer S, Zhen R;
XX	
DR	WPI; 2002-139930/18.
XX	
PT	New cyanobacterial nucleic acid fragments encoding acetylhydroxyacid
PT	synthase (AHA5) or phytoene desaturase (PDS), useful for conferring
PT	herbicide, fungicide or insecticide resistance, and for identifying AHA5
PT	or PDS inhibitors.
XX	
PS	Claim 4; Page 39; 70pp; English.
XX	
CC	The invention relates to cyanobacterial nucleic acid fragments encoding
CC	acetylacetyl synthase (AHA5) large and small subunits
CC	or mutant phytoene desaturase (PDS). Polynucleotides of the invention are
CC	useful for controlling plant traits via nuclear or plasmome
CC	transformation, e.g. for conferring herbicide, fungicide or insecticide
CC	resistance. Particularly, they are useful in identifying novel PDS and
CC	AHA5 inhibitors and in plant transformations for conferring resistance
CC	and cross-resistance to certain bleaching herbicides and AHA5-inhibiting

CC herbicides. The present sequence is *Synechocystis* sp. strain PCC 6803
XX AHS small subunit gene, 11VN
CC
SQ Sequence 565 BP; 141 A; 138 C; 161 G; 125 T; 0 U; 0 Other;

Query Match 100.0%; Score 565; DB 6; Length 565;
Best Local Similarity 100.0%; Pred. No. 1.2e-165;
Matches 565; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGGAATTTTACCCCAATGGCCACCGGATGCTTCTTGGCCCAATGAACACACC 60
DB 1 GTGGAATTTTACCCCAATGGCCACCGGATGCTTCTTGGCCCAATGAACACACC 60
QY 61 CTCTCTGTTTAAAGTGAAGATGAAGCCGAGTGTAAACCCGATTCGCGATTTGGC 120
DB 61 CTCTCTGTTTAAAGTGAAGATGAAGCCGAGTGTAAACCCGATTCGCGATTTGGC 120
QY 121 CGCCGTGTTTAACTTGAAGCTTGGCGGTGGCGGAGGAGGAGGAGTTC 160
DB 121 CGCCGTGTTTAACTTGAAGCTTGGCGGTGGCGGAGGAGGAGGAGTTC 160
QY 181 CGCATCACCATGCTGCTGCGGAGATGAACACCATGCAACATGACCAAGCACTC 240
DB 181 CGCATCACCATGCTGCTGCGGAGATGAACACCATGCAACATGACCAAGCACTC 240
QY 241 TACCAAGTTGTTAAAGTAAAGTACAGACATCAACCAACTCCCTGTGTGAAAGG 300
DB 241 TACCAAGTTGTTAAAGTAAAGTACAGACATCAACCAACTCCCTGTGTGAAAGG 300
QY 301 GAATGATCTGAGGAAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 360
DB 301 GAATGATCTGAGGAAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 360
QY 361 GCCCAAGTATTCGCGGCGCGGATGTGATATCTCCGAGACACCGTACCATGAAATG 420
DB 361 GCCCAAGTATTCGCGGCGCGGATGTGATATCTCCGAGACACCGTACCATGAAATG 420
QY 421 TGGGGAGACCGGATTAATGTTAGCAATCTCCAGATGTCGCAAGTGGCAATTAAGA 480
DB 421 TGGGGAGACCGGATTAATGTTAGCAATCTCCAGATGTCGCAAGTGGCAATTAAGA 480
QY 481 GGTGCTCGAAGCGGCAAAATGCTTGTGTCGGGAAATCCGGCTCAATACGAAATCT 540
DB 481 GGTGCTCGAAGCGGCAAAATGCTTGTGTCGGGAAATCCGGCTCAATACGAAATCT 540
QY 541 GAAATCCTCGAATCCAAATGTTTAA 565
DB 541 GAAATCCTCGAATCCAAATGTTTAA 565

RESULT 2

ABA98820
ID ABA98820 standard; DNA; 1721 BP.

XX ABA98820;

XX 18-JUN-2002 (first entry)

DE Soybean AHS small subunit (clone sdc2c.pk001.b10) coding sequence.

KM AHS; Acetolactate synthase; acetylhydroxy synthase;
KM amino acid biosynthesis; plant; herbicide; AHS small subunit;
KM AHS holozyme; crop protection chemical; expressed sequence tag; EST;
ss.

XX Glycine max.

XX Key Location/Qualifiers

FT CDS 88..1433

FT /*tag= a

FT /product= "AHS small subunit"

FT /EC_number= "4.1.3.18"

XX

PN US200104939-A1.

XX 22-NOV-2001.

XX 08-DEC-2000; 2000US-00732618.

XX 04-JAN-2000; 2000US-0174437P.

XX (ABELL) ABELL L M.

XX (FALCO) FALCO S C.

XX (FAMO) FAMODU O O.

PI Abell LM, Falco SC, Famodu OO;

XX WPI; 2002-113395/15.

XX P-PSDB; ABB08416.

PT New small subunits of plant acetolactate synthase and nucleic acids
encoding them, useful in screening for novel crop protection chemicals or
potential herbicidal compounds based upon holozyme inhibition.

XX Claim 7; Page 37; 57pp; English.

The invention relates to new isolated polypeptides, which comprises an
acetolactate synthase (AHS) small subunit. Acetolactate synthase is the
first committed step in branched chain amino acid biosynthesis in plants
and bacteria. The polypeptides are useful in aiding the discovery of new
herbicides that inhibit plant acetolactate synthase activity. The
polypeptides are useful for preparing plant AHS holozyme, which are
useful in screening for potential herbicidal compounds based upon
holozyme inhibition. The AHS holozyme is also useful in screening for
novel crop protection chemicals. Particularly, the polypeptides can be
used as a target to facilitate design and/or identification of inhibitors
of those enzymes that may be useful as herbicides. Thus, the polypeptides
could be appropriate for new herbicide discovery and design. The
polynucleotides are also useful for producing the polypeptides and for
screening for new herbicides. The current sequence represents the soybean
AHS small subunit (clone sdc2c.pk001.b10) coding sequence

XX Sequence 1721 BP; 472 A; 401 C; 391 G; 457 T; 0 U; 0 Other;
Query Match 25.0%; Score 141.2; DB 6; Length 1721;
Best Local Similarity 58.2%; Pred. No. 3e-33;
Matches 285; Conservative 0; Mismatches 203; Indels 2; Gaps 2;

QY 54 AACACCCCTCTCTGTTTAAAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 113
DB 1005 AACACGTTATCCAGCTTGTGAAGATGCTCTGGAATTTTAAACATTTTACAGAGT 1064
QY 114 ATTGCCCCCGCGTGTGTTTAACTGAGAGCTTGGCGGTGGGATCGCGGAAACAGGGGGA 173
DB 1065 TTTGCTGAAGAGGCTTAACATTCAGAGTTTACTGTAGAACATGACAGAGTTGAAGG 1124
QY 174 CGTTTCCGATCACATGCTGTGCTCCGGGGATGAACACCATCGAACATGACCAA 233
DB 1125 ACTTCTCGACTTAACCTGTGTGCTCGGGACATGATGATTAATGCAAGTTGGTGA 1184
QY 234 GCACTCTAAGTGTGTTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGT 293
DB 1185 GCACTCTAAGTGTGTTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGT 1244
QY 294 GGAAGGAGTATGCTGTGTAAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGT 353
DB 1245 TGAGGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1304
QY 354 TGAGGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 413
DB 1305 TGATTTGCTAAGCTTTTCCGGGCTAAAGCTGTGATATCTGATCAACATTAACCT 1364
QY 414 CGA-ATGTTGGGAGACCGGATTAATGTTAGCAATCTCTCCAGATGTT-GGCCAAAGTTGG 471
DB 1365 GGAGCTTACTGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1424

Qy 472 CATTAAGAGTGGCTGCAGACGGCAAAATTCCTTGGTGGGGAATCCGGGTCAATAC 531
 Db 1425 CATTTGTGAGTGGGACCACTGGGGCAATTCCTTAAGTCCGAGTCCGGGTGACTC 1484
 Qy 532 GGAAATATCTG 541
 Db 1485 CAAGTACTTG 1494

RESULT 3

AA199682_33/c
 Continuation (34 of 45) of AA199682 from base 3300001 (Mycobacterium tuberculosis strain
 WP Sequence Split into 45 fragments Locus AA199682 Accession AA199682

Fragment Name	Begin	End
WP AA199682_00	1	110000
WP AA199682_01	100001	210000
WP AA199682_02	200001	310000
WP AA199682_03	300001	410000
WP AA199682_04	400001	510000
WP AA199682_05	500001	610000
WP AA199682_06	600001	710000
WP AA199682_07	700001	810000
WP AA199682_08	800001	910000
WP AA199682_09	900001	1010000
WP AA199682_10	1000001	1110000
WP AA199682_11	1100001	1210000
WP AA199682_12	1200001	1310000
WP AA199682_13	1300001	1410000
WP AA199682_14	1400001	1510000
WP AA199682_15	1500001	1610000
WP AA199682_16	1600001	1710000
WP AA199682_17	1700001	1810000
WP AA199682_18	1800001	1910000
WP AA199682_19	1900001	2010000
WP AA199682_20	2000001	2110000
WP AA199682_21	2100001	2210000
WP AA199682_22	2200001	2310000
WP AA199682_23	2300001	2410000
WP AA199682_24	2400001	2510000
WP AA199682_25	2500001	2610000
WP AA199682_26	2600001	2710000
WP AA199682_27	2700001	2810000
WP AA199682_28	2800001	2910000
WP AA199682_29	2900001	3010000
WP AA199682_30	3000001	3110000
WP AA199682_31	3100001	3210000
WP AA199682_32	3200001	3310000
WP AA199682_33	3300001	3410000
WP AA199682_34	3400001	3510000
WP AA199682_35	3500001	3610000
WP AA199682_36	3600001	3710000
WP AA199682_37	3700001	3810000
WP AA199682_38	3800001	3910000
WP AA199682_39	3900001	4010000
WP AA199682_40	4000001	4110000
WP AA199682_41	4100001	4210000
WP AA199682_42	4200001	4310000
WP AA199682_43	4300001	4410000
WP AA199682_44	4400001	4411529

Query Match 24.5%; Score 138.4; DB 4; Length 110000;
 Best Local Similarity 57.1%; Pred. No. 1.4e-31;
 Matches 290; Conservative 0; Mismatches 216; Indels 2; Gaps 2;

Qy 22 CACCGCGCATGCCCTCTTTGGCCCGCCGTAAGACACACCTCTGTTTGAAGTGAAGT 81
 Db 61145 CACCGAAGCCACGCGCTGATGATGAGCCCGAAGACGACAGTGTGCTGTTGATGAGAC 61086
 Qy 82 GAAGCCGAGTGTACTAACCGCATTCGCGACTATTGGCCGCGGTGTTTAACATTGAG 141
 Db 61085 AAGCCCGGCTGTGCGGAGTGGCGGCGCTGTTTCCCGCGCGGTTTCAACATCGAG 61026

Db 61025 TCCTTGGCGGTGGTGGCCACCGAGTGCAGAGACAGTACAGTACAGTACGCTGCTCC 60966
 Qy 202 GGGGATGAGAACACCATCGAACACTGACCAAGCAACTTACAGTTGGTTAACTGTAAT 261
 Db 60965 GCCGAGGACACTCCGCTCGAGCAGATCACCAACAGCTCAACAGCTGATCAAGTATC 60906
 Qy 262 AAGTACAGGACATCAACCGAACTCCCTGTGTGGAAGGAAATGATGCTGGGAGAGT 321
 Db 60905 AAGTGTGTCAGACGACGACGACGACTCGGTCTACGGGAATGGCGCTCATCAAGTTC 60846
 Qy 322 AGCGCCATGCCCCCTTAACCGAGCGGAAGTATGAGCTAGCCAGGTAATCCGGGCCGC 381
 Db 60845 CAGCCGACCGCCGACGCGCAGCCGACCAAGTATGAAAGCGGTGATCTGTTCCGCCAAC 60786
 Qy 382 ATTGTGATATCTCCGAGACACCGTCAACATGAAATGAGTGGGACCCG-GGTAAATG 440
 Db 60785 GTGATTCAGCTATCTCCGAGTATTCAGCTGAGGCGCACCGGTAAACCGCGCAAGTTA 60726
 Qy 441 GTAGCAATCCTCCAGATGTTGGCCAAGTT-GGCATTAAAGAGGTGCTCGAACGGGCAAA 499
 Db 60725 GAGGCCCTGTGGGGGTGTGGAGCGGTTGCGTATTCGGAATCCGCCAATCCGGAATG 60666
 Qy 500 ATTGCTTTGGTGGCGGGAATCCGGCGTCA 527
 Db 60665 GTGTGCTGTCCCGCGGTCCGCGCGCA 60638

RESULT 4

AA199683_33/c
 Continuation (34 of 44) of AA199683 from base 3300001 (Mycobacterium tuberculosis strain
 WP Sequence Split into 44 fragments Locus AA199683 Accession AA199683

Fragment Name	Begin	End
WP AA199683_00	1	110000
WP AA199683_01	100001	210000
WP AA199683_02	200001	310000
WP AA199683_03	300001	410000
WP AA199683_04	400001	510000
WP AA199683_05	500001	610000
WP AA199683_06	600001	710000
WP AA199683_07	700001	810000
WP AA199683_08	800001	910000
WP AA199683_09	900001	1010000
WP AA199683_10	1000001	1110000
WP AA199683_11	1100001	1210000
WP AA199683_12	1200001	1310000
WP AA199683_13	1300001	1410000
WP AA199683_14	1400001	1510000
WP AA199683_15	1500001	1610000
WP AA199683_16	1600001	1710000
WP AA199683_17	1700001	1810000
WP AA199683_18	1800001	1910000
WP AA199683_19	1900001	2010000
WP AA199683_20	2000001	2110000
WP AA199683_21	2100001	2210000
WP AA199683_22	2200001	2310000
WP AA199683_23	2300001	2410000
WP AA199683_24	2400001	2510000
WP AA199683_25	2500001	2610000
WP AA199683_26	2600001	2710000
WP AA199683_27	2700001	2810000
WP AA199683_28	2800001	2910000
WP AA199683_29	2900001	3010000
WP AA199683_30	3000001	3110000
WP AA199683_31	3100001	3210000
WP AA199683_32	3200001	3310000
WP AA199683_33	3300001	3410000
WP AA199683_34	3400001	3510000
WP AA199683_35	3500001	3610000
WP AA199683_36	3600001	3710000
WP AA199683_37	3700001	3810000
WP AA199683_38	3800001	3910000

XX WO200277183-A2.
 XX 03-OCT-2002.
 XX 21-MAR-2002; 2002WO-US009107.
 XX 21-MAR-2001; 2001US-00815242.
 XX 06-SEP-2001; 2001US-00948993.
 XX 25-OCT-2001; 2001US-0342923P.
 XX 08-FEB-2002; 2002US-0007851.
 XX 06-MAR-2002; 2002US-0366992P.
 XX (ELIT-1) ELITRA PHARM INC.
 XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
 XX Wall D, Trawick JD, Carr GJ, Yamamoto R, Foreyth RA, Xu HH;
 XX MPI; 2003-029926/02.
 XX P-PSDB; ABU33819.
 XX
 XX New antisense nucleic acids, useful for identifying proteins or screening
 XX for homologous nucleic acids required for cellular proliferation to
 XX isolate candidate molecules for rational drug discovery programs.
 XX
 XX Claim 14; SEQ ID NO 25559; 1766p; English.
 XX
 XX The invention relates to an isolated nucleic acid comprising any one of
 XX the 6213 antisense sequences given in the specification where expression
 XX of the nucleic acid inhibits proliferation of a cell. Also included are:
 XX (1) a vector comprising a promoter operably linked to the nucleic acid
 XX encoding a polypeptide whose expression is inhibited by the antisense
 XX nucleic acid; (2) a host cell containing the vector; (3) an isolated
 XX polypeptide or its fragment whose expression is inhibited by the
 XX antisense nucleic acid; (4) an antibody capable of specifically binding
 XX the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 XX proliferation or the activity of a gene in an operon required for
 XX proliferation; (7) identifying a compound that influences the activity of
 XX the gene product or that has an activity against a biological pathway
 XX required for proliferation, or that inhibits cellular proliferation; (8)
 XX identifying a gene required for cellular proliferation or the biological
 XX pathway in which a proliferation-required gene or its gene product lies
 XX or a gene on which the test compound that inhibits proliferation of an
 XX organism acts; (9) manufacturing an antibiotic; (10) profiling a
 XX compound's activity; (11) a culture comprising strains in which the gene
 XX product is overexpressed or underexpressed; (12) determining the extent
 XX to which each of the strains is present in a culture or collection of
 XX strains; or (13) identifying the target of a compound that inhibits the
 XX proliferation of an organism. The antisense nucleic acids are useful for
 XX identifying proteins or screening for homologous nucleic acids required
 XX for cellular proliferation to isolate candidate molecules for rational
 XX drug discovery programs, or for screening homologous nucleic acids
 XX required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 XX *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target
 XX prokaryotic essential genes. Note: The sequence data for this patent did
 XX not form part of the printed specification, but was obtained in
 XX electronic format directly from WIPO at
 XX ftp.wipo.int/pub/published_pct_sequences

Query Match 24.0%; Score 135.8; DB 7; Length 504;
 Best Local Similarity 56.0%; Pred. No. 8.3e-32;
 Matches 277; Conservative 0; Mismatches 217; Indels 1; Gaps 1;

QY 41 TGCCCCCATGAAACACACCTCTCTTTTAACTGAGATGAAGCCGAGTCTTAACC 100
 DB 2 TGAATCGGACAGACCAACGCTGTGCTGTGCAAGACAAACCGGAGTGTCTGCCC 61
 QY 101 GCATTGCCGACTATTGCCCCCGCGTGTGTTTAACATTGAGACTTGGCGGTGCG 160
 DB 62 GGGGCGGCGCGCTGTTCAGCGCGGCGGTTCACATCGAATGCTTGGCGTGGTCA 121

QY 161 CGAAGCAGGGGAGCGTTTCCGCATCACCATTGTTGTCGCGGGAGTGAACAACATCG 220
 DB 122 CCGAGCAGAAAGACATGTCTGCGGATGACGATGTGTGTCCGCCAGGACCCCGCTCG 181
 QY 221 AACCACTGACCAAGCACTCTTCAAGTTGTGTTAACTTAATTAAGACATCACCG 280
 DB 182 AGCAGATCAACCAAGACGCTCAACAGCTGATCAACGATCAAGATCTTGAAGG 241
 QY 281 AAATCTCCCTGTGGAAGGGAATGATGTCGTGAAGTGAAGCCCAATGCCCTAAC 340
 DB 242 ACGGCACTCGGTGTGTCGGGGAATTTGGCGTTGATCAAGTGTGCGCCGACCC 301
 QY 341 GAGCGGAAGTATTAAGTCTAGCCAGGTAATTCGCGGCCCGCATTTGTGATATCTCCGAG 400
 DB 302 GCAGTCAAGTCAATCAGCGCGGTAACCTGTTCCGCGCAAGGTATGACGTTTCGCGG 361
 QY 401 ACACCGTCAACATCGA-ATGATGAGGGAACCGCGGTAATGATGCAATCTCCAGATG 459
 DB 362 AGCGCTGACATTCAGGCGCACCGGTGACCCGCGAATATGAGCGGTGTGCGGGTGC 421
 QY 460 TGGCAAGTTGCAATTAAGAGTGGCTCGAACCGGCAAAATGCTTTGTCGGGAATC 519
 DB 422 TGAAGCCCTTCGATTCGATGATGTCCTCAATCGGAGTGTGTGTCGCGGATC 481
 QY 520 CGGCGTCAATACGCA 534
 DB 482 CGCGCGAATCGCA 496

RESULT 8
 ABA98816
 ID ABA98816 standard; DNA; 1813 BP.
 AC ABA98816;
 XX 18-JUN-2002 (first entry)
 DT
 XX
 DE Corn ALS small subunit (clone p0094.csc[172ra] coding sequence.
 XX
 XX ALS: Acetylacate synthase; acetylhydroxy synthase;
 KW amino acid biosynthesis; plant; herbicide; ALS small subunit;
 XX ALS holoenzyme; crop protection chemical; ss.
 XX
 XX Zea mays.
 OS
 XX
 FH Key Location/Qualifiers
 FT CDS 1..1452
 FT /*tag=a
 FT /partial
 FT /product= "ALS small subunit"
 FT /EC number= "4.1.3.18"
 FT /note= "no start codon present"
 XX
 PN US2001044939-A1.
 XX
 XX 22-NOV-2001.
 PD
 XX 08-DEC-2000; 2000US-00732618.
 PF
 XX 04-JAN-2000; 2000US-0174437P.
 PR
 XX
 XX (ABELL) ABELL L M.
 PA (FALCO) FALCO S C.
 PA (FAMO/) FAMODU O O.
 XX
 XX Abell LM, Falco SC, Famodu OO;
 XX MPI; 2002-113395/15.
 DR P-PSDB; ABB08412.
 XX
 PT New small subunits of plant acetylacate synthase and nucleic acids
 PT encoding them, useful in screening for novel crop protection chemicals or
 PT potential herbicidal compounds based upon holoenzyme inhibition.

XX Claim 7, Page 31, 57pp; English.
 XX
 CC The invention relates to new isolated polypeptides, which comprises an
 CC acetylactate synthase (ALS) small subunit. Acetylactate synthase is the
 CC first committed step in branched chain amino acid biosynthesis in plants
 CC and bacteria. The polypeptides are useful in aiding the discovery of new
 CC herbicides that inhibit plant acetylactate synthase activity. The
 CC polypeptides are useful for preparing plant ALS holoenzyme, which are
 CC useful in screening for potential herbicidal compounds based upon
 CC holoenzyme inhibition. The ALS holoenzyme is also useful in screening for
 CC novel crop protection chemicals. Particularly, the polypeptides can be
 CC used as a target to facilitate design and/or identification of inhibitors
 CC of those enzymes that may be useful as herbicides. Thus, the polypeptides
 CC could be appropriate for new herbicide discovery and design. The
 CC polynucleotides are also useful for producing the polypeptides and for
 CC screening for new herbicides. The current sequence represents the corn
 CC ALS small subunit (clone p0094.cssl172a) coding sequence
 CC
 SQ Sequence 1813 BP; 545 A; 370 C; 432 G; 466 T; 0 U; 0 Other;

Query Match 24.0%; Score 135.4; DB 6; Length 1813;
 Best Local Similarity 57.5%; Pred. No. 2e-31;
 Matches 281; Conservative 0; Mismatches 206; Indels 2; Gaps 2;
 DB 54 ACACACCCCTCTGTTTAAAGTGAAGAGCCGAGTGTCTAACCCTGATGCGGACT 113
 939 ACATCTCTATCATCTTCTGTAATGATGTCCTGAGATTTCTCATCTCTGTAACAGGTGT 998
 114 ATTGCCCCCGGTGTTTAACTTGAAGCTTGGCGGTGGCGGAAACGGGAGG 173
 999 ATTCTCCAGAGGGGCTCAATATTCAGAGCTTGTGTCGCCAGCTGAAAGAGG 1058
 114 CGTTCCCGCATCAGCATGCTGTCGCCGGGAGTGAAGAACCATTCGAAACAGTCAGCA 233
 1059 AACTTCTGCATCTACTGCTGTTCTCGAAGCTGATATTCATTCGAAAGCTAATGACA 1118
 234 GCAACTCTACAGTGTGTTAACTTAAAGTACAGACATCAACCGAACTCCCTGTGT 293
 1119 TCAACTGTAAAGCTCATGTGATGTTTATGAAGTTTACAGATTTTACCCACTTACATTTGC 1178
 234 GGAAGGGATGATGCTGCTGAGAGCTGAGGCCAATGCCCTTACCGAGCGGAAATGAT 353
 1179 TGCTGAGAGATTAAGATCATTAAGTGGCGCAATCTACAGCTCAAGAGGATGTCTT 1238
 354 TGAGCTAGCCGAGTATTCGCGGCGCGCATTTGTGATTTCCGAGACCGCTAC-CA 412
 1239 AGATTTGCTCGATTTTGGAGGACAGAAAGTTGACATATCAGACCAACAATTAACAT 1298
 413 TCGAATGCTGGGAGCCCGGTAAATGATGACATCTCCAGATGTTGGCCAAAGT-TGG 471
 1299 ACTGCTCAACCGAGATTTGACAGATGTTAGATTTGCAAAAGATGTCAGCGATGAGG 1358
 472 CATTTAAGAGTGGCTCGAAGCGGCAAAATGCTTGTGGGGGGGAATCCGCGCTCAATAC 531
 1359 CATCTGTAGGTTGACGAGACGCGCGGATTTGCTGCTCGAGAGTCTGAGATTGACTC 1418
 532 GGAATATCT 540
 1419 CAAGTACCT 1427

RESULT 9
 AC26110
 ID AC26110 standard; DNA; 489 BP.
 AC AC26110;
 XX
 DT 19-JUN-2003 (first entry)
 XX Prokaryotic essential gene #7767.
 DE Antisense; ds; prokaryotic essential gene; cell proliferation;
 XX

KW drug design; gene.
 XX Burkholderia mallei.
 OS
 PN NC020027183-A2.
 XX
 PD 03-OCT-2002.
 XX
 PF 21-MAR-2002; 2002WC-US009107.
 XX
 PR 21-MAR-2001; 2001US-00815242.
 PR 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362599P.
 XX
 PA (ELIT-) ELITRA PHARM INC.
 XX
 PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zykkind JW;
 PI Wall D, Trawick JD, Carr GT, Yamamoto R, Forsyth RA, Xu HH;
 XX
 DR WPI; 2003-029926/02.
 DR P-PSDB; AB022240.
 XX
 PT New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 XX
 PS Claim 14; SEQ ID NO 13980; 1766pp; English.

XX The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target
 CC prokaryotic essential genes. Note: The sequence data for this patent did
 CC not form part of the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

Sequence 489 BP; 109 A; 140 C; 158 G; 82 T; 0 U; 0 Other;
 Query Match 23.8%; Score 134.6; DB 7; Length 489;
 Best Local Similarity 59.1%; Pred. No. 1.9e-31;
 Matches 230; Conservative 0; Mismatches 159; Indels 0; Gaps 0;
 DB 49 ATGAACACACCCCTCTGTTTAAAGTGAAGAGCCGAGTGTCTAACCCTGATGCG 108
 1 ATGAGACACATCTTTCGCTGCTGAGAAATGAACCGGCGCTGTGCGCGCTGCTG 60
 109 GAATATTTCGCCCGGTGTTTAACTTGAAGAGCTTGGCGGTGGCGGAGCAAG 168

DT	04-DEC-2000 (first entry)
XX	
XX	N. meningitidis partial DNA sequence gnm_32 SEQ ID NO:32.
DE	
DE	
XX	Neisseria meningitidis; Neisseria gonorrhoeae; genome; immunogenic;
KW	antigen; vaccine; diagnosis; infection; antibacterial; identification.
KW	Meningococcus B; MenB; ds.
XX	

OS Neisseria meningitidis.
XX WO20002430-A2.
XX 20-APR-2000.
XX 08-OCT-1999; 99WO-US023573.
XX 09-OCT-1998; 98US-0103794P.
XX 30-APR-1999; 99US-0132068P.
XX (CHIR) CHIRON CORP.
XX Frazer CM, Hickey E, Peterson J, Tettelin H, Venter JC,
PI Masignani V, Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V,
PI Rapuoli R, Pizza M;
XX WPI, 2000-318079/27.
XX
XX Isolated nucleotide sequences of Neisseria meningitidis which can be used
PT in the diagnosis and treatment of N. meningitidis infection and other
PT Neisserial infections, for example, N.gonorrhoea.
XX
XX Claim 7, Page 589-593; 1760pp; English.
XX
XX The present invention describes methods of obtaining immunogenic proteins
CC from Neisseria genomic sequences. AA81453 to AA82414 represent
CC specifically claimed Neisseria meningitidis genomic DNA sequences;
CC AA81266 to AA81303 and AA825620 to AA825663 represent Neisseria DNA
CC sequences and their corresponding proteins; AA81254 to AA81259 and
CC AA81304 to AA81321 represent PCR primers used in the isolation of
CC Neisseria meningitidis DNA sequences; and AA81322 to AA81452 represent
CC Neisseria meningitidis MenB polynucleotide ORF sequences, which are all
CC used in the exemplification of the present invention. The nucleic acid
CC sequences, protein sequences, and antibodies against them, can be used in
CC the manufacture of a composition. The composition can be used as a
CC medicament (or in the manufacture of a medicament) for treating,
CC preventing or diagnosing infection due to Neisserial bacteria. For
CC example, some of the identified proteins could be components of vaccines
CC against Meningococcus B; against all serotypes; and/or against all
CC pathogenic Neisseriae. Identification of sequences from the bacterium
CC will also facilitate production of biological probes, particularly
CC organism-specific probes. Attempts to make efficacious Meningococcus B
CC vaccines have failed mainly due to antigen tolerance. Multivalent
CC vaccines have also been tried but none have successfully overcome
CC antigenic variability. The provision of further, complete sequences may
CC provide an opportunity to identify secreted or surface exposed proteins
CC that may be presumed targets for the immune system and which are not
CC antigenically variable or at least more conserved than other more
CC variable regions
XX
XX Sequence 13732 BP; 3488 A; 3771 C; 3357 G; 3116 T; 0 U; 0 Other;
SQ
Query Match 23.7%; Score 133.8; DB 3; Length 13732;
Best Local Similarity 60.2%; Pred. No. 1.5e-30;
Matches 222; Conservative 0; Mismatches 147; Indels 0; Gaps 0;

QY 289 TGTGTGAAGGAATTTGATGCTGTGATGAGTGAGCGCCCAATGCCCTAACCGAGCGGAA 348
DB 5621 TTGTGCGAAGCGAACTGATGTTGTGTAAGTCCGTGCGCCGCAAGACCGCGACGAA 5562
QY 349 GTGATGAGCTAGCCGAGGTATTCGGGCCCGCATTTGTGATATCTCCGAAGACACCGTC 408
DB 5561 TTTTACGCTTGAACCAATCTACCGGGGAGAGATCATGACGTAAACGACCGAGCTAT 5502
QY 409 ACCATCGAA 417
DB 5501 ACGATTGAA 5493
RESULT 12
AA81489 1/C
Continuation (2 of 9) of AA81489 from base 100001 (N. meningitidis partial DNA sequenc
WP Sequence Split into 9 fragments LOCUS AA81489 Accession Aa81489
WP Fragment Name Begin End
WP AA81489_0 1 110000
WP AA81489_1 100001 210000
WP AA81489_2 200001 310000
WP AA81489_3 300001 410000
WP AA81489_4 400001 510000
WP AA81489_5 500001 610000
WP AA81489_6 600001 710000
WP AA81489_7 700001 810000
WP AA81489_8 800001 837096
Query Match 23.7%; Score 133.8; DB 3; Length 110000;
Best Local Similarity 60.2%; Pred. No. 3.8e-30;
Matches 222; Conservative 0; Mismatches 147; Indels 0; Gaps 0;
QY 49 ATGAAACACACCTCTCTGTTTAACTGATGAAGACCGGAGTGTAAACCCGATTCGC 108
DB 102597 ATGGACATATCTTATCTGTTGATTTGAAAGACATCAAGTCCGATGACCCCGTGC 102538
QY 109 GGAATTTGCCCGCGGTGTTTAACTGAGAGCTTGCGGTGGGTGCGGAAACAG 168
DB 102537 GGTGTTCTCTCCACGCGATTAATCAATGATCTTGTGGCGGTGCGCGAAGAC 102478
QY 169 GGGGAGCTTCCCGCATCAACATGTGTGTCGCGGGGATGAGAACCAATGAAACATG 228
DB 102477 AAAACCTGTACAGATGACCAATGCTTACCAACGCGGACGAGAGTCAATGAAATT 102418
QY 229 ACCAAGCACTTCAAGATTGTTAACTGATTAATTAAGTACAGACATCACCGAACTGCC 288
DB 102417 ACCAAGCACTTCAATTAATGATTAAGTATTAAGTGTGATTTGAATGAAAGCCGT 102358
QY 289 TGTGTGAAGGAATTTGATGCTGTGATGAGTGAGCGCCCAATGCCCTAACCGAGCGGAA 348
DB 102357 TTGTGCAAGCGAACTGATGTTGTGTAAGTCCGTGCGCCGCAAGACCGCGAGAA 102298
QY 349 GTGATGAGCTAGCCGAGGTATTCGGGCCCGCATTTGTGATATCTCCGAAGACACCGTC 408
DB 102297 TTTTACGCTTGAACCAATCTACCGGGGAGAGATCATGACGTAAACGACCGAGCTAT 102238
QY 409 ACCATCGAA 417
DB 102237 ACGATTGAA 102229
RESULT 13
AA81489 2/C
Continuation (3 of 9) of AA81489 from base 200001 (N. meningitidis partial DNA sequenc
WP Sequence Split into 9 fragments LOCUS AA81489 Accession Aa81489
WP Fragment Name Begin End
WP AA81489_0 1 110000
WP AA81489_1 100001 210000
WP AA81489_2 200001 310000
WP AA81489_3 300001 410000
WP AA81489_4 400001 510000
WP AA81489_5 500001 610000
WP AA81489_6 600001 710000

WP AAA81489.7 700001 810000
 WP AAA81489_8 800001 837096
 Query Match 23.7%; Score 133.8; DB 3; Length 110000;
 Best Local Similarity 60.2%; Pred. No. 3.8e-30;
 Matches 222; Conservative 0; Mismatches 147; Indels 0; Gaps 0;
 QY 49 ATGAAACACACCCCTCTGTTTGTAGTGAAGATGAGCGGAGTGTACCCCGCATGGCC 108
 DB 2557 ATGCACATATCTTATCTGTTCTGATTGAAAACGATACAGTGGATGAGCCGGTGTGTC 2538
 QY 109 GGACTATTTGCCCCCGGTGTTTAACTTGAAGAGCTTGGCGTGGCGGTGGCGGAACAG 168
 DB 2537 GGTGTGTTCTCTGACGCGGATTAACATTCATCTTTGGCGGTGGCGGACGGAAGAC 2478
 QY 169 GGGGACGTTCCCGCATCACCATGCTGTGGTGGCGGGGATGAGAACACCATCGAACACATG 228
 DB 2477 AAAACCTGTACGAGATACCATGCTTACCCAGCGGACGACGCTCATCGAACCAATT 2418
 QY 229 ACCAAGCAACTCTACAACTGTTAACTTAACTTAAAGTACAGACATCACCGAAACTGCC 288
 DB 2417 ACCAAGCAACTCAATTAATGATTGAAGTGAATTAAGTGTGCTGATTTGAATGAAGCGGT 2358
 QY 289 TGTGTGGAAGGAAATGATGCTGTGAAGGTGAGCGCCCAATGCCCTTAACCGAGCGGAA 348
 DB 2357 TTTGTGGAACGCGAACTGATGTTGTAAGTCCGTCCGCGCGCAAAAGCCGCGACGAA 2298
 QY 349 GTGATTGAGCTAGCCGACGATATCCGGCCCGCATTTGTGATATCTCGAAGACACCGTC 408
 DB 2297 TTTTACGCTTGACCGAAATCTACCGGGGACGATCATCGATACGTAACCGACGCTAT 2238
 QY 409 ACCATCGAA 417
 DB 2237 ACGATTGAA 2229
 RESULT 14
 AAF21611/c
 ID AAF21611 standard; DNA; 349980 BP.
 AC AAF21611;
 XX
 DT 13-MAR-2001 (first entry)
 XX
 DE Neisseria meningitidis B nucleotide sequence SEQ ID NO:112.
 XX
 KM Neisseria meningitidis; Neisseria gonorrhoeae; immunogenic; vaccine;
 KM diagnosis; antigen; detection; infection; gene therapy; antibacterial;
 KM ds.
 XX
 OS Neisseria meningitidis.
 XX
 PN WO200066791-A1.
 XX
 PD 09-NOV-2000.
 XX
 PF 08-MAR-2000; 2000WO-US005928.
 XX
 PR 30-APR-1999; 99US-0132068P.
 PR 08-OCT-1999; 99WO-US023573.
 PR 28-FEB-2000; 2000GB-00004695.
 XX
 PA (CHIR) CHIRON CORP.
 PA (GENO-) INST GENOMIC RES.
 XX
 PI Pizsa M, Hickey E, Petersen J, Tertelin H, Venter JC,
 PI Maizgani V, Galeotti C, Mora M, Ratti G, Scarselli M, Scarsiato V,
 PI Rappuoli R, Frazer CW, Grandi G;
 XX
 DR WPI; 2000-647603/62.
 XX
 PT Neisseria meningitidis B full length genome sequence and open reading
 PT frames are used to detect, treat and prevent Neisserial infections.

XX
 PS Claim 7; Appendix A; 692pp; English.
 XX
 CC The present invention describes the full length genome of Neisseria
 CC meningitidis B (NMB). The sequences in AAF21544 and AAF21607 to AAF21613
 CC represent fragments of the NMB genomic sequence, and the sequence which
 CC long to go in a record on its own it was split into 8 sequences which
 CC overlap each other at the beginning and end of each sequence by 49980 bp
 CC (i.e. the last 49980 bp of AAF21544 are repeated at the beginning of
 CC AAF21607, the last 49980 bp of AAF21607 are repeated at the beginning of
 CC AAF21608, and so on). AAF21545 to AAF21588 encode the Neisseria proteins
 CC given in AAB58550 to AAB58593, and AAF21589 to AAF21606 represent PCR
 CC primers which are used in the exemplification of the present invention.
 CC The NMB genome and fragments from it have antibacterial activity, and can
 CC be used in vaccines and gene therapy. Neisseria nucleic acids, proteins
 CC and/or antibodies which binds to the proteins can be used in compositions
 CC for treating or preventing infection due to Neisserial bacteria or as a
 CC diagnostic reagent for detecting the presence of Neisserial bacteria or
 CC of antibodies raised to Neisserial bacteria. Computers, computer memory,
 CC computer storage medium or computer databases can be used in a search to
 CC identify open reading frames (ORFs) or coding sequences within the NMB
 CC genome. The DNA sequences provide further opportunities to find antigenic
 CC or immunogenic proteins which are more effective in vaccines than the
 CC outer membrane proteins currently used
 XX
 SQ Sequence 349980 BP; 87189 A; 93501 C; 84627 G; 84663 T; 0 U; 0 Other;
 Query Match 23.7%; Score 133.8; DB 3; Length 349980;
 Best Local Similarity 60.2%; Pred. No. 5.3e-30;
 Matches 222; Conservative 0; Mismatches 147; Indels 0; Gaps 0;
 QY 49 ATGAAACACACCCCTCTGTTTGTAGTGAAGATGAGCGGAGTGTACCCCGCATGGCC 108
 DB 137826 ATGCACATATCTTATCTGTTCTGATTGAAAACGATACAGTGGATGAGCCGGTGTGTC 137767
 QY 109 GGACTATTTGCCCCCGGTGTTTAACTTGAAGAGCTTGGCGTGGCGGTGGCGGAACAG 168
 DB 137766 GGTGTGTTCTCTGACGCGGATTAACATTCATCTTTGGCGGTGGCGGACGGAAGAC 137707
 QY 169 GGGGACGTTCCCGCATCACCATGCTGTGGTGGCGGGGATGAGAACACCATGGAACAATG 228
 DB 137706 AAAACCTGTACGAGATACCATCTTACCCAGCGGACGACGAGTGTATGAAACAAATT 137647
 QY 229 ACCAAGCAACTCTACAACTGTTAACTTAAAGTAAAGTACAGACATCACCGAAACTGCC 288
 DB 137646 ACCAAGCAACTCAATTAATGATTGAAGTGAATTAAGTGTGCTGATTTGAATGAAGCGGT 137587
 QY 289 TGTGTGGAAGGAAATGATGCTGTGAAGGTGAGCGCCCAATGCCCTTAACCGAGCGGAA 348
 DB 137586 TTTGTGGAACGCGAATGATGTTGTAAGTCCGTCCGCGCAAAAGACCGGACGAA 137527
 QY 349 GTGATTGAGCTAGCCGAGTATTCGGGGCCGCGCATTTGTGATATCTCGAAGACACCGTC 408
 DB 137526 TTTTACGCTTGACCGAAATCTACCGGGGACGATCATCGATACCGACCGAGCTAT 137467
 QY 409 ACCATCGAA 417
 DB 137466 ACGATTGAA 137458
 RESULT 15
 ABA98814
 ID ABA98814 standard; DNA; 1297 BP.
 AC ABA98814;
 XX
 DT 18-JUN-2002 (first entry)
 XX
 DE Corn ALS small subunit (clone csn3n.pK0112.c11) coding sequence.
 XX
 KM ALS; Acetolactate synthase; acetylhydroxy synthase;
 KM amino acid biosynthesis; plant; herbicide; ALS small subunit;
 KM ALS holoenzyme; crop protection chemical; ss.

QY	322	AGGCGCAATGCGCCCTTAACCCAGGGGAGTGTATTAGCTACGCCAGGTAATTCGCGGGCCCGC	381
Db	3355177	CAMCCCGACGCGCGGACGCGCAGCCCAAGTAGTGTCAAGGGGTGAATCTGTTTTCGCGCCAC	3355118
QY	382	ATTGTGGATATCTCTCCGAAGCACCGGTACCACTCCAAATGTTGGGGGACCGG-GCTTAAATG	440
Db	3355117	GTCATTACAGTATCCCGGAGTCAATTGACGCTCGAGCGCACCGGTTAACCGGGGCAAGTTA	3355058
QY	441	GTACGAATCTCTCCAGATGTTGGCCAAATT-GGCATTTAAAGGTTGGCTCGAACGGGCAA	499
Db	3355057	GAGGCGCTGCTGCGCGGGTGTGGACCGCTCGGTATTCGCGAAATCGCCCAATCCGGAATG	3354938
QY	500	ATTGCTTTGGGCGGGAATCCGGGCTCA	527
Db	3354997	GTCGCGCTGTCCCGCGGTCGCGCGGCA	3354970

```

RESULT 2
US-09-103-840A-1/C
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103.840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
US-09-103-840A-1

```

	Query Match	24.5%	Score 38.4	DB 3	Length 4411529
	Best Local Similarity	57.1%	Pred. No. 2.2e-33		
	Matches 290	Conservative 0	Mismatches 216	Indels 2	Gaps 2
QY	22	CACGCGGATGCGCTTCTTTGCCCCCATGAAACACACCCCTCTGTGTTAGTTGAAGAT	81		
Dh	3361145	CACCGAAGGCGCACCCCTGATGAGCCCGAAGCGCACAGTGTGCGTGTGTGTGCGAAGAC	33610866		
QY	82	GAAGCGGAGTGCCTAACCCCGCATGSCCGGACATTTGCGCGCGGTGTTTAAACATGAG	141		
Dh	3361085	AAAGCCGCGGTGCTGCGCGCGGTGCGCGCTGTGTTCTCCGCGCCGGTTTCAACATGAG	33610266		
QY	142	AGCTTGGCGGTGGGGGTGCGCGGAAACAGGGGACCTTTTCCGCATCACCATGTTGGTCCG	201		
Dh	3361025	TGTTTGGCGGTGGGTGCGCACGAGTGCAGGAGACAGTCAAGATGACCATGTGTGTC	33609666		
QY	202	GGGATGAGAAACACCATCGAACCACTGACCAAGCACTCTTACATGTTGGTTAACTAATT	261		
Dh	3360965	GCCGAGACATCTCGCTGAGCAATATCACCAAGGCTCAACAAGCTGATCAAGTATC	33609066		
QY	262	AAATGACGAGCATCACCGAACTCCCTGTGTGAAAGGGAAATTGATGCTGTGTAAAGT	321		
Dh	3360905	AAAGTGTGAGGACGAGAGACGAGACATCGGTGTCAAGGGAATTGGGCTCATCAAGT	33608466		
QY	322	AGCGCCAAATGCCCCCTTAACCGAGCGGAAGTGAATTGAGTGAAGCCAGGATTCCTGGGCCCCG	381		
Dh	3360845	CAAGCCGACGCGCGGACACCCGACCGCAAGTAGTCAAGACCGGTGAATCTGTTTCGGCCAA	33607866		
QY	382	ATTGTGATATCTTCCGAGACACCGTCACCATCGAATGTGTGGGAGACCG-GGTAAATG	440		
Dh	3360785	GTTATTGACGATTCGCCGAGTCAATTGACGTGTGAGGACACCGGTTAACCGGCAAGTTA	33607266		
QY	441	GTAAGCAATCTCCAGATGTTGGCCAGTT-GCATTTAAAGGTGGCTCGAACGGCGAA	499		

Db	Qy	Db
3360725	500	3360665
GAGGCGCTGTCGGGGGTGTGGAGCCGTTGGGATTCGGCAATCGCCCAATCCGGAAATG	ATTGCTTTGGTGGCGGAATCCGGCGTCA	GTCGCGCTGTCGGCGCGGTCGCGCGGGA
	527	3360638

```

RESULT 3
US-09-252-991A-7249
; Sequence 7249, Application US/09252991A
; Patent No. 6551785
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 7249
; LENGTH: 540
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-7249

```

```

Query Match          20.8%; Score 112.8; DB 4; Length 540;
Best Local Similarity 55.0%; Pred. No. 4,9e-27;
Matches 219; Conservative 0; Mismatches 177; Indels 0; Gaps 0;

QY      48  CATTGAAACACACCCCTCTCTGTTTACTGTGAAGATGAAGCCGAGTGTCTAACCCGCAATTGC 107
DB      48  CATGCGCACATCATATTTCTCCCTGCTGTGGAAGAACGACCAAGGCGCATTTGTCCCGCTGGT 107

QY      108  CGGACTATTTGCCCCCGCCGCGTGTTTTAAACATTGAAGCTTGGCCGATGGGAGT 167
DB      108  CGGCGCTTGTCTCCCAAGCGCAACTAACCAATCGAAAGGCTTACCTGTGGCCGCGCACCGAGA 167

QY      168  GGGGGAGCTTTCCCGGATCATCCATGGTGGTGGCCGGGGAGATGAGAAACCATCGAACACT 227
DB      168  CCGGACCTCTTCCGCTGTGACGCTGACCACTCGTGGGCGACAGATGAGGTGATTCAGACGAT 227

QY      228  GACCAAGCAACTCTTACCAAGTGGTTAACGTAATTAAAGTACAGACATCACCGAACTCC 287
DB      228  CACCAAGAACCTTACAACAAGCTGATCGAAGTGGTCAAGCTGTGATGTTCGGAAAGCGC 287

QY      288  CTGTGTGGAAGGAAATTGATGCTGGGAAAGTGAGAGCCATATGCCCTTACAGGAGCGGA 347
DB      288  CCATATGAGCGCGGAGGTGATGCTGTGTGAAGTCAAGGCCACCGGCGCCACGCGCGCGCA 347

QY      348  AGTGATTTAGCTAGCCCAAGGTATTTCCGGGCGCCGCAATTGTGATATCTTCGAAAGACACCGT 407
DB      348  GGTCAAGCCGACCAACCATATCTTCGATGGGAGATCGTGAAGTCAACGATGAGCGTCTA 407

QY      408  CACCATCGATGTGTGGGGAGCCCGGGTAAATAGTGA 443
DB      408  TACCGTGCAACTGGCGGGTACCAAGCGCAAAACTGGA 443

RESULT 4
US-09-252-991A-7450/C
; Sequence 7450, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A

```

Query Match 20.0%; Score 112.8; DB 4; Length 1722;
Best Local Similarity 55.3%; Pred. No. 8.9e-27;
Matches 219; Conservative 0; Mismatches 177; Indels 0; Gaps 0;

;; CURRENT FILING DATE: 1999-02-18
;; PRIOR APPLICATION NUMBER: US 60/074,788
;; PRIOR FILING DATE: 1998-02-18
;; PRIOR APPLICATION NUMBER: US 60/094,190
;; PRIOR FILING DATE: 1998-07-27
;; NUMBER OF SEQ ID NOS: 33142
;; SEQ ID NO 7450
;; LENGTH: 1722
;; TYPE: DNA
;; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-7450

QY 48 CATGAACACACCCCTCTCTGTTTAAAGATGAAGCCGAGTGAACCCGCAATTGC 107
DB 1654 CATGGAACACATCATTTCCCTGCTGCTGAAAAGAGCCAGCCGATTTGCCCGTGT 1595
QY 108 CGGACTATTGCGCCCGCTGTTTAAATTGAAGCTTGGCGGTGGCGGAAACA 167
DB 1594 CGGCGTGTCTCCCAACCACTACACATCGAAAGCCTGACCGTGGCGCCGACGAGGA 1535
QY 168 GGGGGACGTTTCCGCTATCACCATGATGATGGTGGCGGGAGTGAAGAACACATCGAACT 227
DB 1534 CCGGACCTGTGCGCTGTAGAGCTGACCAACCGTGGGACGATGAGTGCAGCAT 1475
QY 228 GACCAAGCACTCTTCAAGATTGTTTAAAGTAAAGTACAGACATCACCGAACTCC 287
DB 1474 CACCAAGAACTCTCAACAGCTGATGAAAGTGAAGTGTGATCTGTGGAAGAGCGC 1415
QY 288 CTGTGTGAAGAGGAAATTGATGCTGTGAAGTGAAGCCCATGCGCTTACCGAGGGA 347
DB 1414 CCATTCAGAGCGGCTGATGCTGTGAAGTCAAGGCGCAGCGCGCCGCGGCGA 1355
QY 348 AGTATGAGCTAGCCCAAGTATTCGGGCGCGCATTTGTGATATCTCCGAAGACACCGT 407
DB 1354 GGTCAAGGCGCACACCGATTCCTCGTGGGAGATCTGCAAGCTACCAATAGCGTCTA 1295
QY 408 CACCATCGATGTGGGGGACCCCGGTAAATGTA 443
DB 1294 TACCGTGAACCTGGCGGGTACACGAGCAAACTGGA 1259

RESULT 5
US-09-426-568A-1
; Sequence 1, Application US/09426568A
; Patent No. 6348643
; GENERAL INFORMATION:
; APPLICANT: Kakefuda, Genichi
; APPLICANT: Costello, Colleen
; APPLICANT: Sun, Ming
; APPLICANT: Hu, Weiming
; TITLE OF INVENTION: Genes and Vectors for Conferring Herbicide Resistance
; FILE REFERENCE: 008103/195497
; CURRENT APPLICATION NUMBER: US/09/426,568A
; CURRENT FILING DATE: 1999-10-22
; PRIOR APPLICATION NUMBER: 60/106,239
; PRIOR FILING DATE: 1998-10-25
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1673
; TYPE: DNA
; ORGANISM: Arabidopsis sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (42)..(1514)
; OTHER INFORMATION: Mature Peptide
US-09-426-568A-1

Query Match 19.8%; Score 112; DB 4; Length 1673;
Best Local Similarity 54.5%; Pred. No. 1.6e-26;
Matches 266; Conservative 0; Mismatches 220; Indels 2; Gaps 2;

QY 55 CACACCCCTCTGTTTAAAGTGAAGCGAGGAGCTTAACCCGATTCGGGACTA 114
DB 999 CATCTCTATCATGTTGTTTAATGATATTCAGAGGTTCTTAATTTGACTGTGT 1058
QY 115 TTGCCCCCGGTGTTTAAACATTGAGACTTGGCGGTGGCGGAAAGGGGAC 174
DB 1059 TTGCTCGAAGGGAGTACATATCCAGACTTGGCGGTGAGCATGTGAAACCAAGGCG 1118
QY 175 GTTTCGGGATACACATGATGGTGGCGGGGATGAGAACACATCGAACACCAAG 234
DB 1119 ATTTCAGCATTAACAAGTATTAACCTGCAACAGATGAATCGGTACGAATTTGTGCG 1178
QY 235 CAACCTACAGATTGTTTAAAGTAAATTAAGTACAGACATCAACGAACTCCGTGTG 294
DB 1179 CAATTTAAACATCTAGATGTGATGATGATGATGATGATGATGATGATGATGATGAT 1238
QY 295 GAAAGGAATTGATGCTGTGTAAGTGAAGCGCCATGCCCCCTTAACCGAGGAGTAT 354
DB 1239 GAAAGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1298
QY 355 GACTAGCCGAGTATTCGGGCGCCGATTTGTGATATCTCCGAAACACCGTACCAT - 413
DB 1299 GACTTGTAGTATTTTTCAGGGGTAAAGCTGTGACGTATGATACACATTAATCTTG 1358
QY 414 CGAATGATGGGAGACCCGGAATTAATGTGACATCTTCCAGATGTTCG - CCAAGTTGGC 472
DB 1359 CAGCTTAATCGGGATCTAAGCAAGATGTTGACCTGCAAGGTTATTTGAGCCCTATGT 1418
QY 473 ATTAAGAGGTGCTGAAACGGGCAAAATGCTTGTGTGGGGAATCCGCGTCAATAG 532
DB 1419 ATATGAGGTTTCAAAACCGGTGCTGTGATTTGCTGTGATGATGAGGATGACTCC 1478
QY 533 GAATATCT 540
DB 1479 AAGTACT 1486

RESULT 6
US-09-489-039A-5192
; Sequence 5192, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1998-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 5192
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-5192

Query Match 18.7%; Score 105.8; DB 4; Length 525;
Best Local Similarity 54.5%; Pred. No. 9.4e-25;
Matches 212; Conservative 0; Mismatches 177; Indels 0; Gaps 0;

QY 49 ATGAACACACCCCTCTCTGTTTAAAGTGAAGACCGAGTGTAAACCCGATTTGCC 108
DB 34 ATGGCCGGAATATTAATGATTAATGAAAGAAATCGGATGCAATTTGCCGCTTATC 93
QY 109 GACTATTTGCGCGCGGTGTTTAAACATTGAAGCTTGGCGGTGGCGGTGGCGGAAAG 168
DB 94 GGCTCTTTTCCAGGCTGCTAATATAGAAAGCTGATGCTGCGCGCGGACGAGAT 153

QY 169 GGGGACGTTTCCCGCATCAGATGTTGCGGGGGATGAGAACACCATGAACTG 228
; Sequence 1, Application US/09557884
; Patent No. 6506581
Db 154 CCCACTCTCTCAGAAATGACATTCAGACCGTGGCCATGAAAAAGGATGACAAATT 213
; GENERAL INFORMATION:
QY 229 ACCAAGCACTTACAACTGTTTACCTTAATTAAGTACGACATCCGAAACTCC 288
; APPLICANT: Fieschmann et al.
; TITLE OF INVENTION: The Nucleotide sequence of
; the Haemophilus influenzae Rd Genome, Fragments
; Thereof, and Uses Thereof
Db 214 GAAAAGCAATGCAACAACTGTTGATCTCTGCGGGGACAGCGAGCGGGCT 273
QY 289 TGTGTGGAAGGAATGATGCTGTTGAGGTGAGCCGCAATGCCCCCTAACCGAGGAA 348
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESS: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: MD
; COUNTRY: USA
; ZIP: 20850
Db 274 CATGTGAAACGGAAATTAATCTGTTGAGGTGAGCCGACGCGCTATGCGCGGAA 333
QY 349 GTGATTGAGCTAGCCCGAGTAATCCGGCCCGCATTTGTGATATCTCGAAGACAGCTC 408
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3 1/2 inch diskette
; COMPUTER: Dell Pentium
; OPERATING SYSTEM: MS DOS v6.22
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/557,884
; FILING DATE: 25-Apr-2000
; CLASSIFICATION: <Unknown>
Db 334 GTGAAGCCCAACCCGAATCTCTCCGCGACAGATTAATGACGTGAGCGCTTCTATCTAT 393
QY 409 ACCATCGAATGTTGGGGAGCCCGGGTAAA 437
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/476,102
; FILING DATE: JUN-5-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Michelle S. Marks
; REGISTRATION NUMBER: 41,971
; REFERENCE/DOCKET NUMBER: P8186P3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 301-309-8504
; TELEFAX: 301-309-8439
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1830121 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; US-09-557-884-1

RESULT 7
US-09-328-352-3858
; Sequence 3858, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 3858
; LENGTH: 507
; TYPE: DNA
; ORGANISM: Acinetobacter baumannii
US-09-328-352-3858

Query Match 18.5%; Score 104.6; DB 4; Length 507;
Best Local Similarity 54.2%; Pred. No. 2.3e-24;
Matches 212; Conservative 0; Mismatches 179; Indels 0; Gaps 0;
QY 49 ATGAAACACACCCCTCTGTTTATGTAAGATGAGCCGAGTGTAAACCCGATGGC 108
Db 16 ATGAGACACATTTTCTGTTACTGTTGAAAACGAGCTGTGGCTTTCTCTTTGTTG 75
QY 109 GCACTATTGCCCCCGCTGTTTAACTTGAAGCTTGGCGGTGGCGGAGCAG 168
Db 76 GCGTATTCACTCAACGTAACTACAACTTGAACACTTAATGTGACCAACCGAAGAT 135
QY 169 GGGGACGTTTCCCGCATCAGATGTTGCGGGGGATGAGAACACCATGAACTG 228
Db 136 CCGACATTATGCGCTTGAAGCTTAACGACTTATGCGCATGACCTAAGATGAACTAAT 195
QY 229 ACCAAGCACTTCAAGTGTGTTAAGTAATTAAGTAAAGACATCAGCAAACTCC 288
Db 196 ACGAAACAGCTCAACAGCTTGTGAGTGTGTAAGTAAAGTATGATTATCAGAAAGTTCA 255
QY 289 TGTGTGGAAGGAATGATGCTGTTGAGGTGAGCCGCAATGCCCCCTAACCGAGGAA 348
Db 256 CACATGAGCGTGAATCTATGCTAATCAAGTAAACATTAAGTGTTCACGCGCTGAA 315
QY 349 GTGATTGAGCTAGCCCGAGTAATCCGGCCCGCATTTGTGATATCTCGAAGACAGCTC 408
Db 316 ATTAAGACCACTGGGGAATTTTCCGTGCGCAATTTGACGTGACACCAAGACCTAT 375
QY 409 ACCATCGAATGTTGGGGAGCCCGGGTAAA 437
Db 376 ACCATCGAATGACGAGGACAACTGAAAAAT 406

RESULT 8

US-09-557-884-1/c
; Sequence 1, Application US/09557884
; Patent No. 6506581
; GENERAL INFORMATION:
; APPLICANT: Fieschmann et al.
; TITLE OF INVENTION: The Nucleotide sequence of
; the Haemophilus influenzae Rd Genome, Fragments
; Thereof, and Uses Thereof
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESS: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: MD
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3 1/2 inch diskette
; COMPUTER: Dell Pentium
; OPERATING SYSTEM: MS DOS v6.22
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/557,884
; FILING DATE: 25-Apr-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/476,102
; FILING DATE: JUN-5-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Michelle S. Marks
; REGISTRATION NUMBER: 41,971
; REFERENCE/DOCKET NUMBER: P8186P3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 301-309-8504
; TELEFAX: 301-309-8439
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1830121 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; US-09-557-884-1
Query Match 17.7%; Score 100.2; DB 4; Length 1830121;
Best Local Similarity 54.1%; Pred. No. 4.5e-21;
Matches 204; Conservative 0; Mismatches 173; Indels 0; Gaps 0;
QY 62 TCTCTGTTTATGTTAAGATGAGCCGAGTGTAAACCCGATTCGCGACTATTGCCC 121
Db 1651189 TATCTGTTTACTCGAAATGATCGGAGCATTTGCGTGTGTTTATTTCTCTC 1651130
QY 122 GCGGTGTTTAACTTGAAGCTTGGCGGTGGGGTGGCGGAAACAGGGGACGTTCCC 181
Db 1651129 AAGTGCCTTAAATTAAGTGTGACGTTGACCAACGAGATGATCAACGCTTCTC 1651070
QY 182 GCATCACATGTTGTTGCGGGGGATGAGAACACCATGAACTGAACCTGCAAGCACTCT 241
Db 1651069 GATGACATTTGAACCTGTGGCGCATGCTCAGGATTTGACCAATGAAACAACTTC 1651010
QY 242 ACAAGTTGTTAAGTAAATTAAGTACAGGACATCCGAACTCCCTGTGTGAAAGGG 301
Db 1651009 ATTAATTAAGTATGTTTAAAGTGTGTTAATTTAAGTACAGAACTCTTAAGACAG 1650950
QY 302 AATTGATGCTGTTGAGTGAAGTGAAGCCCAATGCCCCCTAACCGAGGGAAGTGTAGCTAG 361
Db 1650949 AATTTGTTTACGAAAGTAAAGACAGTAGTTATCTCGGATGAAATTTAAAGATTAG 1650890
QY 362 CCCAGATATCCGGGCGCATTTGTGATATCTCGAAGACACCGTACCATGGAATGTG 421
Db 1650889 CAGATATTTTCCGCGCTCAAAATTTGATGATGATGACCAAAAGTCTTACACAAATTA 1650830
QY 422 GGGGAGACCCGGGTAAAA 438


```

? Patent No. 6632636
? GENERAL INFORMATION:
? APPLICANT: Lagace, Robert, E.
? APPLICANT: Paterson, Chandra
? APPLICANT: Berg, Kim, L.
? TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF MORAXELLA CATARRHALIS GENOME
? FILE REFERENCE: PM-0008-4 US
? CURRENT APPLICATION NUMBER: US/09/596,002
? CURRENT FILING DATE: 2000-06-16
? PRIOR APPLICATION NUMBER: 60/140,121
? PRIOR FILING DATE: 1999-06-18
? NUMBER OF SEQ ID NOS: 41
? SOFTWARE: PERL Program
? SEQ ID NO 35
? LENGTH: 96109
? TYPE: DNA
? ORGANISM: M. catarrhalis
? FEATURE:
? NAME/KEY: misc feature
? OTHER INFORMATION: Incyte template ID No. 6632636 35
? PUBLICATION INFORMATION:
? US-09-596-002-35

```

Query Match	16.0%	Score 90.2	DB 4	Length 96109	
Best Local Similarity	53.9%	Pred. No. 1,86-18			
Matches 256	Conservative	0	Mismatches 288	Indels 11	Gaps 3
QY	52	AAACACACCCTCTCTGTTTATGTAAAGTAAACCCGAGTCTAACCCGCAATTGCCGGA	111		
Db	4168	AAACATCTGATTTCGTATTATATGAAAAACAAACAGATTCGGTGTGTTGTTCGGT	4227		
QY	112	CTATTGCCCCCGCGTGTATTAACATATAGAGCTTGCGGTGGGGTTCGCGCAACAGGG	171		
Db	4228	TATATTTCACAGCGGGCATTAATATGATACGTCAATATACAGCCAGATATACCA	4287		
QY	172	GACGTTTCCCGATACCATGTGTGTCGGGGGATGAAACCATCGAACACTGACC	231		
Db	4288	ACGATCTCAGGCTACATTTGCAACACATTCACACATGACAAATTCGAAACAAATCAC	4347		
QY	232	AAGCAACTCTACAGTGTGTTAACGTATTTAAATGACAGACATCACCGA-----A	282		
Db	4348	AAAGCACTGCAATTAAGTATGCAAGTGTGTTAAATGTCCTGTCGAGCTGTCAGGC	4407		
QY	283	ACTCCCTGTGGAAAGGAAATGATGTGTGTAAGGTGAGCGCCCAATGCCCTTACCGA	342		
Db	4408	GGTCATCATGTTGAACGAGAGTGTATGCTCATTTAAGTGGTGCACACGGGTTCAAACT	4467		
QY	343	GGGAAGATTTAGACTAGCCCAAGATTCCTCGGGCCCGCATTTGATATATTCGCCAAGC	402		
Db	4468	GAAAGAAATAGCGTATATGCGGATATATTTGTCGCCAAATATGTCATGTCAGCACAAAC	4527		
QY	403	ACCGTCACCA-TCGATGTGTGGGAGACCCGGGTAAATGTGTACCAATCTCTCAGATG-TT	460		
Db	4528	TGTATACAAATTTTAACTTACTGTGTAGTGTGCGCAAAATGTGATGATGTTTATCGATGTGAT	4587		
QY	461	GGCCAAAGTTGCGATTAAGAGATGGCTCGAACGGGCAAAATTCCTTTGGTGGGG	515		
Db	4588	GGTCGTGACCGTATCCTTGAAGTGTGATGCGTCAAGTGTGATTTGTATCGCTGTG	4642		

RESULT 12
US-09-790-988-1/c
; Sequence 1, Application US/09790988
; Patent No. 6632935
; GENERAL INFORMATION:
; APPLICANT: SHIGENOBU, SHUJI
; APPLICANT: MATSUNABE, HIDEMI
; APPLICANT: HATTORI, MASAHITA
; APPLICANT: SAKAKI, YOSHIYUKI
; TITLE OF INVENTION: GENEOME DNA OF BACTERIAL SYMBIONT OF APHIDS
; FILE REFERENCE: 081356/0159
; CURRENT APPLICATION NUMBER: US/09/790,988
; CURRENT FILING DATE: 2001-02-23

```

: PRIOR APPLICATION NUMBER: JP2000-107160
: PRIOR FILING DATE: 2000-04-07
: NUMBER OF SEQ. ID NOS: 7
: SOFTWARE: PatentIn Ver. 2.1
: SEQ. ID NO. 1
: LENGTH: 640681
: TYPE: DNA
: ORGANISM: Buchnera sp.
: US-09-790-988-1

```

Query Match	13.8%	Score 78	DB 4	Length 640681
Best Local Similarity	50.0%	Pred. No. 4.8e-14		
Matches 195	Conservative	0	Mismatches 195	Indels 0
QY	49	ATCAAAACACACCTCTCTGTTTTAGTTGAATGAAGCCGAGTCTTACCCGCAATTGCC	108	
Db	248814	ATCCGAAGATTTTATCTATTTTATGGAAAATGAATCAGGTGCATCATCAAGAGTATATA	248755	
QY	109	GGACATTTATCCCGCGCGTGTTTTACATTTAGAGAGCTGGGCGTGGGGGTGGCGGAACAG	168	
Db	248754	GGACTTTTTCACAAAGAGATATATATTAAGAACTTATACGCTTAGCACTTCTGAAGAT	248695	
QY	169	GGGAGCGTTTCCCGCATCACCATGTGTGTCGCGGAGTAGAACAACATCGAACACTG	228	
Db	248694	CTACTTTATCTAAATGACATATACAAACAGTAGAGAAATGAAAAGCTATGGAACAAAT	248635	
QY	229	ACCAAGCACTTCACAGTGGTTAGCTATTTAAAGTAGCAGCACTCACCGAACTCCC	288	
Db	248634	GAAAAACAATCCATTAATTAATAGATTTATTAAGATGTGTACAAATTGACAACTCTCT	248575	
QY	289	TGTGTGAAAGGAATTAATGCTGTGTAAGGTGACGCGCAATGCCCTTAACGAGCGGA	348	
Db	248574	CATATGAACGGGAATCGTATTATTAAATTTGAGATTAACATCATTAAGAAGAGAT	248515	
QY	349	GTGATTGAGCTAGCCCAAGTATTCGGGCGCCGCAATTGTGATATCTCCGAAGCACCCGTC	408	
Db	248514	ATTAACAATATTACTGAAGTATTTGAGGCGCAATTGTGACATTAATCATCTGCAACATAT	248455	
QY	409	ACCATGATGGTGGGGAACCCGGGTAAA	438	
Db	248454	ATATTCGAATGTCTGGTACTACAAAAA	248425	

```

1      RESULT 13
2      US-08-961-527-8
3      Sequence 8, Application US/08961527
4      Patent No. 6420135
5      GENERAL INFORMATION:
6      APPLICANT: Charles Kunsch
7      TITLE OF INVENTION: Streptococcus pneumoniae Polymnucleotides and Sequences
8      NUMBER OF SEQUENCES: 391
9      CORRESPONDENCE ADDRESS:
10     ADDRESSEE: Human Genome Sciences, Inc.
11     STREET: 9410 Key West Avenue
12     CITY: Rockville
13     STATE: Maryland
14     COUNTRY: USA
15     ZIP: 20850
16     COMPUTER READABLE FORM:
17     MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
18     COMPUTER: HP Vectra 486/33
19     OPERATING SYSTEM: MSDOS version 6.2
20     SOFTWARE: ASCII Text
21     CURRENT APPLICATION DATA:
22     APPLICATION NUMBER: US/08/961,527
23     FILING DATE:
24     CLASSIFICATION: 424
25     PRIOR APPLICATION DATA:
26     APPLICATION NUMBER:
27     FILING DATE:
28     ATTORNEY/AGENT INFORMATION:
29     NAME: Brookes, A. Anders
30     REGISTRATION NUMBER: 36,373

```

REFERENCE/DOCKET NUMBER: PB340P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 6211 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-961-527-8

Query Match 13.2%; Score 74.8; DB 4; Length 6211;
Best Local Similarity 50.9%; Pred. No. 4,8e-14;
Matches 230; Conservative 0; Mismatches 217; Indels 5; Gaps 2;

```
QY 72 AGTGAAGATGAAGCCGAGAGTGTACCGGATTCGGGACTATTTGCCCGCGTGGTTT 131
Db 1802 ACTACAAATGCTTAGAGAGTCTCATGCTTTACAGGTGTCCTATCTGTCGTACAGT 1861
QY 132 TAACTTGAAGCTTGGCGGTGGGTCCGCGAACAAGGCGGACGTTTCCCGCATCACCAT 191
Db 1862 TAACTTGAAGCATCTGTTGGAGCAACAAGATCCGAATGATCGGTATCACTAT 1921
QY 192 GGTGTGCCGGGGATG--AGAACACATCGAACAACTGACCAAGCACTCTCAAGTT 248
Db 1922 TATTATTGATGTGCTTCTCTCATGTAAGTGGAGCAATCATCAAAACGCTCAATCGTCA 1981
QY 249 GATTACCGTAAATTAAGTACAGACATCACCGAACTCCCTGTGTGAAAGGGAATGAT 308
Db 1982 GATTGATGTGATTCGATTCGAGATATACAGACAGAGCTCATTTGGAGCCGAGGTGAT 2041
QY 309 GCTGTGAAGTGAAGCCCAATGCCCTTAACCGAGCCGAAGTGAATGAGCTAGCCCAAGT 368
Db 2042 TTTGTGTAGATGTCTAGCGCCGACGCTGAGAAAGAGCTGAGATTTTAAAGATATTCAAC 2101
QY 369 ATTCGGGCGCCGATTTGATATCTCGAAGACACCGTCAACATCGAATGTTGGGGGAC 428
Db 2102 TTTCCTGCAACAGTAGTAGAGCTAGCCCAAGCTCGATTTCCATTCAGATGACGGGAAA 2161
QY 429 CCGGTAATATGTA--GCAATCTCTCCAGATGTTGGCCCAAGTGGCATTTAAAGAGTGC 486
Db 2162 TGCAAAAGAGCGAAGCCCTATTGCGAGTCATTCGCCCAACGATTCGCAATATGTC 2221
QY 487 TCGAAGCGGCAAAATTCCTTTGTCGCGGGAA 518
Db 2222 TCGAAGCGGTGCACTGAGATTTACCGCGGAT 2253
```

RESULT 14
US-08-916-421B-1
Sequence 1, Application US/08916421B
Patent No. 6503729
GENERAL INFORMATION:
APPLICANT: Built et al.
TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanococcus
Patent No. 6503729
FILE REFERENCE: PB275
CURRENT APPLICATION NUMBER: US/08/916,421B
CURRENT FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: US 60/024,428
PRIOR FILING DATE: 1996-08-22
NUMBER OF SEQ ID NOS: 3
SOFTWARE: PatentIn version 3.1
SEQ ID NO 1
LENGTH: 1664976
TYPE: DNA
ORGANISM: Methanococcus jannaschii
FEATURE:
NAME/KEY: misc_feature
LOCATION: (28252)..(28222)
OTHER INFORMATION: n equals a, t, c, or g

```
NAME/KEY: misc_feature
LOCATION: (28257)..(28258)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (84773)..(84773)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (84808)..(84808)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (84812)..(84812)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (98120)..(98120)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (98159)..(98159)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (98239)..(98239)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (98266)..(98266)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (98343)..(98343)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (103988)..(103988)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (148948)..(148948)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (163385)..(163385)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (191989)..(191989)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (191995)..(191995)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (231980)..(231980)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (234187)..(234187)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (234220)..(234220)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (234814)..(234814)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (309398)..(309398)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (309418)..(309418)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (312837)..(312837)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (312993)..(312993)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (319226)..(319226)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (559167)..(559167)
OTHER INFORMATION: n equals a, t, c, or g
```

```

LOCATION: (559241)..(559241)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (600992)..(600992)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (622708)..(622708)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (657081)..(657081)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (657203)..(657203)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (674435)..(674435)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (682442)..(682442)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (713652)..(713652)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (741684)..(741684)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (779455)..(779455)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (796766)..(796766)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (855539)..(855539)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (871619)..(871619)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1084830)..(1084830)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1096846)..(1096846)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1119881)..(1119881)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1130881)..(1130881)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1310988)..(1310988)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1313224)..(1313224)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1349473)..(1349473)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1349491)..(1349491)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1470091)..(1470091)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1569020)..(1569020)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1602912)..(1602912)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1603734)..(1603734)

```

```

OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1637998)..(1637998)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1664855)..(1664855)
OTHER INFORMATION: n equals a, t, c, or g
US-08-916-421B-1

```

Query Match 12.4%; Score 70; DB 4; Length 1664976;
 Best Local Similarity 51.1%; Pred. No. 3.3e-11;
 Matches 240; Conservative 0; Mismatches 225; Indels 5; Gaps 3;

```

QY 55 CACACCCCTCTGTTTGTGTAAGATGAACCCGAGCTTAACCCGATTCGGGACTA 114
DB 1307211 CATGTACTCTCGCGTGTGTTTAAATTAACAGAGATTTCAGAGATTTCAGGTTA 1307270
QY 115 TTTCGCCCGCGTGTGTTTAACTGAGAGCTTGCGGTGCGGCGAACAAG3GAGAC 174
DB 1307271 TTACAGAGAGAGGGTTTAAATTTCAAGTATTAACAGATTAACAGAGATTCACAA 1307330
QY 175 GTTTCGCCGATCACATGCTGTGTCGGGGATGAGAACACCATCGAACCACTGACCAAG 234
DB 1307331 ATTTCAGAGAGTACGATGTTGTTAAATGAGATGATTAAGATTAAGAGAGTTATCAA 1307390
QY 235 CAACCTCAAGATTGTTAGTAACTAATTAAGTACAGACATCACCAACTCCCTGTGTG 294
DB 1307391 CAACCTCAAGATTGTTAGTAACTAATTAAGTAACTGATTAAGAGAGAGAGAGAGAG 1307450
QY 295 GAAAGGAATTGATCTGCTGTAAGTGACG---CAATGCCCTTAACCGAGCGAAGTG 351
DB 1307451 CAGAGAGAGCTCTGTTTAAATTAAGATTAATCAACCAAGAGAGTGAAGAAATCAAGTT 1307510
QY 352 ATTCAGTACGACAGTATTCGCGGCGCGCATGTGATATCTCGAAGACACCGTCAAC 411
DB 1307511 ATTCAATTAACAGATTAATTAAGAGAGAAATGTTAGATTTAACTCCAGAAATCTTTAATT 1307570
QY 412 ATTCGA-ATGCTGGGGAGACCCGCTGAATATGTTAGCAATCTCCAGAT-GTTGGCCAAATT 469
DB 1307571 GTAGAGATTAACGTGATGAGAGATTAATTAACCACTTATTGACTTAAGTAAACCACTTA 1307630
QY 470 GGCATTAAAGAGGTGCTCGAAGCGGCAAAATTGCTTGTGGCGGGAATC 519
DB 1307631 GGAATTAAGAAATGCGCAAGACTGAAATTAATCTCCTTAAGCAAGGGAATC 1307680

```

```

RESULT 15
US-09-426-568A-3
; Sequence 3, Application US/09426568A
; Patent No. 6348643
; GENERAL INFORMATION:
; APPLICANT: Kakufuda, Genichi
; APPLICANT: Costello, Colleen
; APPLICANT: Sun, Ming
; APPLICANT: Hu, Weiming
; TITLE OF INVENTION: Genes and Vectors for Confering Herbicide Resistance
; FILE REFERENCE: 008103/195497
; CURRENT FILING DATE: US/09/426,568A
; PRIOR APPLICATION NUMBER: 1999-10-22
; PRIOR FILING DATE: 1998-10-29
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 4895
; TYPE: DNA
; ORGANISM: Arabidopsis sp.
; FEATURE:
; NAME/KEY: Promoter
; LOCATION: (1)..(757)
; OTHER INFORMATION: Promoter Region
; NAME/KEY: misc_feature

```

```

; LOCATION: (717)
; OTHER INFORMATION: Transcriptional Starting Point
; NAME/KEY: misc_signal
; LOCATION: (758)..(760)
; OTHER INFORMATION: Start Codon
; NAME/KEY: misc_signal
; LOCATION: (4737)..(4739)
; OTHER INFORMATION: Stop Codon
; OTHER INFORMATION: n at position 694 can be a, c, g, or t
US-09-426-568A-3

```

```

Query Match      10.9%; Score 61.6; DB 4; Length 4895;
Best Local Similarity 58.7%; Pred. No. 8.8e-10;
Matches 128; Conservative 0; Mismatches 84; Indels 6; Gaps 1;

```

```

QY      49 ATGAACAACACCCCTCTCTGTTTAACTGAAGTGAAGCCGGAGTGCTAACCCGATTGCC 108
DB      1088 AGAAGCAACACATTTTCAATTTGTGTGAGCGAAGCGAATGATTATAGATTGCA 1147
QY      109 GGAATATTTGGCCCGCGTGGTTTAACTGAGAGCTTGGCGTGGGTGGCGGAACAG 168
DB      1148 GAGGTGTTGCAAGAGAGATACATATTTGAGAGTCTTGTGTTGCTG-----AAC 1201
QY      169 GGGGACGTTTCCCGCATCACCATGCTGGTGGCGGGGATGAGAACACCATGAAACAAGT 228
DB      1202 AGAGCAAGGCTCTATTCACCATAGTTGTCTGTGSACTGAAGGTAATTCAAGAGGTC 1261
QY      229 ACCAAGCAACTCTACAGTTGGTTAAGTTAAAGT 266
DB      1262 ATGAGCAACTCCAGAGCTGCTTAATGTCTAAAGT 1299

```

Search completed: July 25, 2004, 12:54:57
 Job time : 54.0987 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 25, 2004, 07:28:18 / Search time 272.203 Seconds

(without alignments)
10146.437 Million cell updates/sec

Title: US-09-893-033-17

Perfect score: 565
Sequence: 1 gtaagatttaccacatg.....ccctgaatcccaagtttag 565

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 3216467 seqs, 244419694 residues

Total number of hits satisfying chosen parameters: 6432934

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database: Published Applications NA:*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/ECT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/ECTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US08A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
19: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	141.2	25.0	1721	9	US-09-732-618-17
2	137	24.2	504	13	US-10-282-122A-26262
3	137	24.2	507	13	US-10-282-122A-28638
4	135.8	24.0	504	13	US-10-282-122A-25559
5	135.4	24.0	1813	9	US-09-732-618-9
6	134.6	23.8	489	13	US-10-282-122A-13980
7	133.8	23.7	492	13	US-10-282-122A-29826
8	132.8	23.5	1297	9	US-09-732-618-5
9	132.2	23.4	489	13	US-10-282-122A-29150
10	129.6	22.9	525	15	US-10-156-761-7719
11	129.6	22.9	9025608	15	US-09-732-618-15
12	129	22.8	1435	9	US-10-282-122A-15104
13	127.6	22.5	489	13	US-10-282-122A-13405
14	126.6	22.4	489	13	US-10-282-122A-13405

15	126	22.3	525	13	US-10-282-122A-11953	Sequence 11953, A
16	132.8	21.7	1525	13	US-10-424-599-23993	Sequence 23993, A
17	117.2	20.7	510	13	US-10-282-122A-27732	Sequence 27732, A
18	116.6	20.6	1561	17	US-10-437-963-3796	Sequence 3796, Ap
19	114	20.2	2256646	17	US-10-470-565-1	Sequence 1, Appl
20	112	19.8	960	9	US-09-770-445-311	Sequence 311, Appl
21	112	19.8	1673	9	US-09-997-900-1	Sequence 7946, Ap
22	111.8	19.8	492	9	US-09-815-242-7946	Sequence 30523, A
23	111.8	19.8	1861	13	US-10-282-122A-30523	Sequence 23, Appl
24	110	19.5	1861	9	US-09-732-618-23	Sequence 3458, Ap
25	109.4	19.4	516	9	US-09-738-626-3458	Sequence 1, Appl
26	109.4	19.4	3309400	9	US-09-738-626-1	Sequence 5844, Ap
27	108.4	19.2	1464	17	US-10-437-963-5844	Sequence 7233, Ap
28	105.8	18.7	519	13	US-10-282-122A-7233	Sequence 22330, A
29	105.8	18.7	519	13	US-10-282-122A-33330	Sequence 2950, Ap
30	105.6	18.7	556	17	US-10-021-323-2950	Sequence 17529, A
31	104.4	18.5	522	13	US-10-282-122A-17529	Sequence 31295, A
32	104.2	18.4	489	9	US-09-815-242-3987	Sequence 7645, Ap
33	104.2	18.4	522	9	US-09-815-242-7645	Sequence 30933, A
34	103.8	18.4	489	13	US-10-282-122A-31295	Sequence 41304, A
35	103	18.2	489	13	US-10-282-122A-9034	Sequence 7159, Ap
36	101.8	18.0	492	13	US-10-282-122A-30933	Sequence 1, Appl
37	100.6	17.8	485	13	US-10-282-122A-41304	Sequence 12917, A
38	100.2	17.7	492	9	US-09-815-242-7159	Sequence 22364, A
39	100.2	17.7	492	13	US-10-282-122A-22364	Sequence 1, Appl
40	100.2	17.7	1830121	15	US-10-329-670-1	Sequence 12917, A
41	100.2	17.7	1830121	16	US-10-329-670-1	Sequence 1, Appl
42	99.6	17.6	2130	13	US-10-424-599-129217	Sequence 41728, A
43	97.6	17.3	495	13	US-10-282-122A-41728	Sequence 33628, A
44	97.4	17.2	489	13	US-10-282-122A-33628	Sequence 38994, A
45	95.8	17.0	492	13	US-10-282-122A-38994	

ALIGNMENTS

RESULT 1					
US-09-732-618-17					
Sequence 17, Application US/09732618					
Patent No. US2001004939A1					
GENERAL INFORMATION:					
APPLICANT: Falcu, Lynn					
APPLICANT: Abell, Carl					
APPLICANT: Falcu, Omolayo O.					
TITLES OF INVENTION: Small Subunit of Plant Acetolactate Synthase					
FILE REFERENCE: BB1435 US NA					
CURRENT APPLICATION NUMBER: US/09/732,618					
CURRENT FILING DATE: 2000-12-08					
PRIOR APPLICATION NUMBER: 60/174,437					
PRIOR FILING DATE: 2000-01-04					
NUMBER OF SEQ ID NOS: 43					
SOFTWARE: Microsoft Office 97					
SEQ ID NO 17					
LENGTH: 1721					
TYPE: DNA					
ORGANISM: Glycine max					
US-09-732-618-17					
Query Match					
Best Local Similarity 58.2%; Pred. No. 4, 2e-37;					
Matches 285; Conservative 0; Mismatches 203; Indels 2; Gaps 2;					
QY	54	ACACACCCCTCTGTTTGTAGTAAAGTAAAGCCGAGTCTACCCGACT	113		
DB	1005	ACACACCTTATCATGCTGTGACATGCTCTGAGATCTTAAATTTGACAGAGT	1064		
QY	114	ATTGCCCCCGGCTTTTAACTTGAAGCTTGGCGGTGCGCGAAGAGGGA	173		
DB	1065	TTTGTCTAAGAAAGGCGTATTAACATTCAGAGTTAGCTTGAAGCATGCAAGAGTTAAG	1124		
QY	174	CGTTTCCGATACCATGCTGTCGGCGGATGGAACACATCACTGAACACTGACCA	233		
DB	1125	ACTTTCGACTTACACTGTGTCTCTGGACAGATGATCAATTAGCAAGTTGTGTGA	1184		

ORGANISM: Mycobacterium bovis
US-10-282-122A-26262
Query Match 24.2%; Score 137; DB 13; Length 504;
Best Local Similarity 57.7%; Pred. No. 6.3e-36;
Matches 282; Conservative 0; Mismatches 205; Indels 2; Gaps 2;
QY 234 GCACTCTACAGTTGGTTACGTAAATTAAGTACAGACATCACCGAAATCCCTGTGT 293
DB 1185 GCACTCTATAGTAGTAGAGGTACATGAGGTTCTGATATATACCCACTTGCATTGGC 1244
QY 234 GGAAGGGAATGATGCTGCGTGAAGGTGAGGCCAAGCCCTTAACCGAGGAATGAT 353
DB 1245 TGACCGAATTTGATGCTAAATTAAGATTGCTGTAAGATGCTGCACGCGTATGCTT 1304
QY 354 TGACCTAGCCAGGATATCCGGGCCCGCATGTTGATATCTCCGAAACACCGTACCAT 413
DB 1305 TGATATGCTAGCATTTTCCGGGCTAAAGCTGTGATGTATCTGATACACATTAATCTT 1364
QY 414 CGA-ATGCTGGGGGACCCGGGTAAATGTTAGCATCTCTCAATGTT-GGCCAAGTTGG 471
DB 1365 GGAGCTTACTGAGATTTTGAACAAGATGTTGCAATTCAGAGATTTGTAACCATATGG 1424
QY 472 CATTAAAGAGGTGCTCGAACCGGCAAAATGCTTTGTCGGGAAATCCGCGTCAATAC 531
DB 1425 CATTGTGAGGTGCGACGAACTGGGCGAATTGCTCTAGTCCGCGAGTCCGCTGTGACTC 1484
QY 532 GGAATATCTG 541
DB 1485 CAAGTACTTG 1494
RESULT 2
US-10-282-122A-26262
Sequence 26262, Application US/10282122A
GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/282.122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PAM.
NUMBER OF SEQ IDS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 26262
LENGTH: 504
TYPE: DNA

ORGANISM: Mycobacterium bovis
US-10-282-122A-26262
Query Match 24.2%; Score 137; DB 13; Length 504;
Best Local Similarity 57.7%; Pred. No. 6.3e-36;
Matches 282; Conservative 0; Mismatches 205; Indels 2; Gaps 2;
QY 41 TCCCCCATTAACACACCCCTCTCTGTTTAAAGTGAAGAGACCGGAGTGTAAACC 100
DB 2 TGAAGCCGAAAGCGACAGCTTGTGCGGTGCGAAGACAGCCGCGCTGCGCC 61
QY 101 GCATTGCCGACTATTTTCCCGCCGCTGTTTAAATTGAGAGCTTGGCGGTGGCGG 160
DB 62 GGGTGGCGCGCTGTTCTCCGGGCGCGTTTCAACATGAGTCTGTGGCGGTGGCGCA 121
QY 161 CGGAACAGGGGACCTTTCCCGCATCAACATGCTGTGTCGGGGAGTGAACACCATCG 220
DB 122 CCGATGCAAGACAGGTCAGGATGACATCGGTGCTCCGCGAGACACTCCGCTCG 181
QY 221 AACACGTGACCAAGCACTCTACAAAGTTGTTAAGTAAATTAAGTACAGGACATCACCG 280
DB 182 AGCATATCACCAAGCATCTCAACAGCTGATCAACGTATCATCAAGATCGTGAAGACG 241
QY 281 AAATCTCCTGTGTGGAAGAGGATGATGCTGTGAGGTGAGGCGCCATGCTTAAAC 340
DB 242 ACAGACACTCGGTGTCAGCGGAATTGGCGCTCATCAAGTCAAGCCGACCCGCGAGCC 301
QY 341 GAGCGGAAGTATGATAGCTAGCCAGGATTCGGGCGCCGATGTTGATATCTCCGAAG 400
DB 302 GAGCGCAAGTATGATCAAGCGGTGAATCTGTTTCCGCCCAAGTATGATGATACCTCCG 361
QY 401 ACACCGTACCATTCGATGATGAGGAGACCG-GETAAATGTTAGCAATCTCCAGATGT 459
DB 362 AGTCATTGACCGTTCAGGCGCACCGGTAAACCGCGCAAGTTAGAGCCCTGCGGGGTGT 421
QY 460 TGGCCAAATT-GGCAATTAAAGGTGTGCTCAACCGGGAATAATTGCTTTGTTGGCGAAT 518
DB 422 TGAAGCCGTTGCGTATTCGCAAAATGCGCAATCGGAATGATGTGCTGTCCCGGCTG 481
QY 519 CCGGCTCA 527
DB 482 CGCGCGCA 490
RESULT 3
US-10-282-122A-28638
Sequence 28638, Application US/10282122A
GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/282.122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06

PRIOR APPLICATION NUMBER: 60/230,347
 PRIOR FILING DATE: 2000-09-09
 PRIOR APPLICATION NUMBER: 60/242,578
 PRIOR FILING DATE: 2000-10-23
 PRIOR APPLICATION NUMBER: 60/253,625
 PRIOR FILING DATE: 2000-11-27
 PRIOR APPLICATION NUMBER: 60/257,931
 PRIOR FILING DATE: 2000-12-22
 PRIOR APPLICATION NUMBER: 60/267,636
 PRIOR FILING DATE: 2001-02-09
 PRIOR APPLICATION NUMBER: 60/269,308
 PRIOR FILING DATE: 2001-02-16
 Remaining Prior Application data removed - See File Wrapper or PALM.
 NUMBER OF SEQ ID NOS: 78614
 SOFTWARE: PatentIn version 3.1
 SEQ ID NO: 28638
 LENGTH: 507
 TYPE: DNA
 ORGANISM: Mycobacterium tuberculosis
 US-10-282-122A-28638

Query Match 24.2%; Score 137; DB 13; Length 507;
 Best Local Similarity 57.7%; Pred. No. 6,3e-36;
 Matches 282; Conservative 0; Mismatches 205; Indels 2; Gaps 2;

41 TGGCCCCCATGAAACACACCTCTCTGTTTGAAGATGAACCGGAGTCTAACCC 100
 2 TGAAGCCCAAGACGACACGTTGTGCGTTGTGAAGACAGCCCGCTGCTGCGC 61
 101 GCATTGCCGACTATTGTCGCCCGCTGTTTAACTTGAAGCTGCGGTGCGGTGG 160
 62 GGGTGGCGGCGCTGTTCTCCGCGCGGTTTCAATGAGTGGTGGCGGTGGTCCA 121
 161 CGGAACAGGGGACCTTTCCCGCATCACCATGTGTGTGCGGGGATGAAACACCATCG 220
 122 CCGAGTGAAGGACAGGTGACGATGACCATGCTGTCTCGCCGAGACACTCGCTCG 181
 221 AACAACTGACCAAGCAACTCTTCAAGTTGTTTAACTTAAATTAAGTACAGACATCACCG 280
 182 AGCAATACCAAGACAGCTCAACAAGCTGATCAAGTATCTCGAGACAGACG 241
 281 AAATCTCTGTGTGAAGAGGATGATGTGTGTGAAGGTGACCGCAATGCCCTTAAAC 340
 242 AGAAGACATCTGCTGTGACGGAATGTGGCTCATCAAGCTCCAAAGCCGACGCGACGC 301
 341 GAGCGGAAGTGAATGAGTGAAGCCAGGTATTCCGGGCGCGCATTTGTGATATCTCCAG 400
 302 GAGCGCAAGTGAATGAAGCGGTGAATCTGTTTCGCGCAAGTGAATGACGTATCTCCGG 361
 401 ACACCGTCAACATGCAATGTGTGGGAGCCCG-GGTAAATGATGCAATCTCTCAAGATG 459
 362 AGTCAATGACCGTCAAGGACCAAGGTAAACGGGCAAGTGAAGGCTGCTGCGGGGT 421
 460 TGGCCAAAGT-GGCATTAAAGAGTGTGCTCAACCGGCAAAATGCTTGTGTGGGAAT 518
 422 TGAAGCCGTTGCGTATGCGCAAAATGCCAATCGGAATGATGTGCTGTCGCGCGGT 481
 519 CGGCGTGA 527
 482 CGGCGGCA 490

RESULT 4
 US-10-282-122A-25559
 Sequence 25559, Application US/10282122A
 Publication No. US20040029129A1
 GENERAL INFORMATION:
 APPLICANT: Wang, Liangsu
 APPLICANT: Zamudio, Carlos
 APPLICANT: Malone, Cheryl
 APPLICANT: Haselbeck, Robert
 APPLICANT: Ohlsen, Kari
 APPLICANT: Zyskind, Judith

APPLICANT: Wall, Daniel
 APPLICANT: Trawick, John
 APPLICANT: Carr, Grant
 APPLICANT: Yamamoto, Robert
 APPLICANT: Forsyth, R.
 APPLICANT: Xu, H.
 TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
 FILE REFERENCE: ELITRA.034A
 CURRENT APPLICATION NUMBER: US/10/282,122A
 CURRENT FILING DATE: 2003-02-20
 PRIOR APPLICATION NUMBER: 60/191,078
 PRIOR FILING DATE: 2000-03-21
 PRIOR APPLICATION NUMBER: 60/206,848
 PRIOR FILING DATE: 2000-05-23
 PRIOR APPLICATION NUMBER: 60/207,727
 PRIOR FILING DATE: 2000-05-26
 PRIOR APPLICATION NUMBER: 60/230,335
 PRIOR FILING DATE: 2000-09-06
 PRIOR APPLICATION NUMBER: 60/230,347
 PRIOR FILING DATE: 2000-09-09
 PRIOR APPLICATION NUMBER: 60/242,578
 PRIOR FILING DATE: 2000-10-23
 PRIOR APPLICATION NUMBER: 60/253,625
 PRIOR FILING DATE: 2000-11-27
 PRIOR APPLICATION NUMBER: 60/257,931
 PRIOR FILING DATE: 2000-12-22
 PRIOR APPLICATION NUMBER: 60/267,636
 PRIOR FILING DATE: 2001-02-09
 PRIOR APPLICATION NUMBER: 60/269,308
 PRIOR FILING DATE: 2001-02-16
 Remaining Prior Application data removed - See File Wrapper or PALM.
 NUMBER OF SEQ ID NOS: 78614
 SOFTWARE: PatentIn version 3.1
 SEQ ID NO: 25559
 LENGTH: 504
 TYPE: DNA
 ORGANISM: Mycobacterium avium
 US-10-282-122A-25559

Query Match 24.0%; Score 135.8; DB 13; Length 504;
 Best Local Similarity 56.0%; Pred. No. 1.6e-35;
 Matches 277; Conservative 0; Mismatches 217; Indels 1; Gaps 1;

41 TGGCCCCCATGAAACACACCTCTCTGTTTGAAGATGAACCGGAGTCTAACCC 100
 2 TGAAGCCCAAGACGACACGCTGTGCGTTGTGAAGACAGCCCGCTGCTGCGC 61
 101 GCATTGCCGACTATTGTCGCCCGCTGTTTAACTTGAAGCTGCGGTGCGGTGG 160
 62 GGGTGGCGGCGCTGTTCTCCGCGCGGTTTCAATGAGTGGTGGCGGTGGTCCA 121
 161 CGGAACAGGGGACCTTTCCCGCATCACCATGTGTGTGCGGGGATGAAACACCATCG 220
 122 CCGAGTGAAGGACAGGTGACGATGACCATGCTGTCTCGCCGAGACACTCGCTCG 181
 221 AACAACTGACCAAGCAACTCTTCAAGTTGTTTAACTTAAATTAAGTACAGACATCACCG 280
 182 AGCAATACCAAGACAGCTCAACAAGCTGATCAAGTATCTCGAGACAGACG 241
 281 AAATCTCTGTGTGAAGAGGATGATGTGTGTGAAGGTGACCGCAATGCCCTTAAAC 340
 242 AGGCAACTGCGTGTGCGGGAATGTGCTTAAGGTGCGCGCCGACGCGGACCC 301
 341 GAGCGGAAGTGAATGAGTGAAGCCAGGTATTCCGGGCGCGCATTTGTGATATCTCCAG 400
 302 GAGTGAAGTGAATGAAGCGGTGAATCTGTTTCGCGCAAGTGAATGACGTATCTCCGG 361
 401 ACACCGTCAACATGCAATGTGTGGGAGCCCG-GGTAAATGATGCAATCTCTCAAGATG 459
 362 AGGCGCTGAAGTGAAGGACCAAGGTAAACGGGCAAGTGAAGGCTGCTGCGGGTGC 421
 460 TGGCCAAAGTGGATTAAGAGGTGCTCAACCGGCAAAATGCTTGTGTGGGAATC 519

Db 422 TGGAGCCCTTCGATCCGATGAGATGCTCAATCGAGAGTGTGTGTTGCCGCGTTC 481
QY 520 CGGCGTCAATACGGA 534
Db 482 CGGCGGAATCGGCA 496

RESULT 5
US-09-732-618-9
; Sequence 9, Application US/09732618
; Patent No. US20010044939A1
; GENERAL INFORMATION:
; APPLICANT: Abell, Lynn
; APPLICANT: Falco, Carl
; APPLICANT: Famodu, Omolayo O.
; TITLE OF INVENTION: Small Subunit of Plant Acetolactate Synthase
; FILE REFERENCE: B81435 US NA
; CURRENT APPLICATION NUMBER: US/09/732,618
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/174,437
; PRIOR FILING DATE: 2000-01-04
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 9
; LENGTH: 1813
; TYPE: DNA
; ORGANISM: Zea mays
US-09-732-618-9

Query Match 24.0%; Score 135.4; DB 9; Length 1813;
Best Local Similarity 57.5%; Pred. No. 4.2e-35;
Matches 281; Conservative 0; Mismatches 206; Indels 2; Gaps 2;

QY 54 ACACACCTCTCTGTTTATTGTTGAAGATGAAGCCGAGTGTAAACCCGATTCGCGACT 113
Db 939 ACATACCTCTATCAATCTCTGTTGAATGATGCTCCGAGTCTCATTTGTAAACAGTGT 998
QY 114 ATTCCCGCCGCTGTTTAACTGATGAGAGCTTGGCGGTGGCGGAGCAAGCGGGA 173
Db 999 ATTCCCGAGAGGGGCTTCAATTTCAAGAGTCTGCTGTTGCCCACTGAAAAGAGG 1058
QY 174 GCTTCCCGCATCACCATGTGTGTGCGCGGAGATGAACACCATCGAACAAGTACCA 223
Db 1059 AACCTTCGATCACTACTGTTGTTCCGAGTGAATGAAATCCATTCGCAAGTAGTACA 1118
QY 234 GCAACTCTACAGTTGTTAAGTAACTAAGTACAGGACATCCGAAACTCCCGTGT 293
Db 1119 TCACTGTACAGCTCATGATGTTTATGAGTTCAAGATTTTACCACTTACCATTGTC 1178
QY 294 GAAAAGGAAATGATGTGTGTAAGGTGAGGCGCAATGCCCTAACCGAGGGAAGTAT 353
Db 1179 TGCTAGAGAGTTAATGATCATMAAGTCCGCGCAATGCTACAGCTGAAAGGATGCTT 1238
QY 354 TGAGTACCCAGGATTTCCGCGCGCGGATGTGATATCTCCGAAAGACACCGCTAC-CA 412
Db 1239 AGATATTGCTCAGATTTTGAAGGACAGAAAGTTGACATACACCAACAAATTACCT 1298
QY 413 TCGAATGTGTGGGAGCCCGGCTAAATATGTAGCATCTCCAGATTTGGCCAGT-TGG 471
Db 1239 ACTGTCTACCGGAGACATTGACAGATGTAGATGAAAGATGCTAGACAGTATGG 1358
QY 472 CATTAAGAGTGTGCTGCAACGCGCAAAATTTGCTTTGTGTGGGAATCCGCGCTCAATAC 531
Db 1359 CATGTGTAGGTGACGACGAGCGCGGATGCTGTCTCGAAGTCTGGAAGTGACTC 1418
QY 532 GGAATATCT 540
Db 1419 CAATGACT 1427

RESULT 6
US-10-282-122A-13980
; Sequence 13980, Application US/10282122A

; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; PRIOR FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,076
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patent version 3.1
; SEQ ID NO 13980
; LENGTH: 489
; TYPE: DNA
; ORGANISM: Burkholderia mallei
US-10-282-122A-13980

Query Match 23.8%; Score 134.6; DB 13; Length 489;
Best Local Similarity 59.1%; Pred. No. 4.1e-35;
Matches 230; Conservative 0; Mismatches 159; Indels 0; Gaps 0;

QY 49 ATGAACACACCTCTCTGTTTATTGTTAAGTGAAGCCGAGTGTAAACCCGATTCG 108
Db 1 ATGAGACACATATTTCCGCTGCTGTAAGAAATGAACCGGCGGCTGTCCGGGTGTC 60
QY 109 GACATATTGCCCCCGCTGTGTTTAACTTGAAGCTTGGCGGTGGGCTGGCGGAACG 168
Db 61 GGTCTGTTCTCCGCAACCGGCTGACCAATCGMACTTGAAGTGGCGCCGACCGAAGAC 120
QY 169 GGGAGCGTTTCCCGATCCATCGATGAGTGGCGGCGGAGTGAACACCATCGAACAATG 228
Db 121 CAGTGGCTGTCCGCTGACATCTCTCCATCGGCTCCGACGATGATGAACAGATC 180
QY 229 ACCAAGCACTCTACAGTGTGTTAAGTAAATTAAGTACAGACATCAACGAAATCTCC 288
Db 181 ACGAAGCATCTGAAACCGCTGATCGAGGTGTGAAGTGTGACCTGACCGACGCTCA 240
QY 289 TGTGTGAAGAGGAAATGATGCTGTGAAGTGAAGCGGCAATGCCCTTAACGAGCGGAA 348
Db 241 CACATGAGCGGAGCTGATGCTCATCAAGTGAAGGAGTGTGGCAAGAGCGGAAAGA 300
QY 349 GTGATTGAGTACCCAGATATTCGCGGCGCGCATTTGATGATCTCGAAGAACCCGCTC 408

Db 301 ATGAAGCGATGCGGACATTTTCGGCGCCGCATCATCGAGTGACGAAAGACCTAC 360
 Qy 409 ACCATCGAATGTGGGGACCCGGGTAA 437
 Db 361 ACGATCGAATTGACGGGGCGAGCGACAA 389

RESULT 7

US-10-282-122A-29826
 ; Sequence 29826, Application US/10282122A
 ; Publication No. US20040029129A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, Liangsu
 ; APPLICANT: Zamudio, Carlos
 ; APPLICANT: Malone, Cheryl
 ; APPLICANT: Haselbeck, Robert
 ; APPLICANT: Ohlsen, Kari
 ; APPLICANT: Zyskind, Judith
 ; APPLICANT: Wall, Daniel
 ; APPLICANT: Trawick, John
 ; APPLICANT: Carr, Grant
 ; APPLICANT: Yamamoto, Robert
 ; APPLICANT: Forsyth, R.
 ; APPLICANT: Xu, H.

; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
 ; FILE REFERENCE: ELITRA.034A
 ; CURRENT FILING DATE: 2003-02-20
 ; PRIOR APPLICATION NUMBER: 60/191,078
 ; PRIOR FILING DATE: 2000-03-21
 ; PRIOR APPLICATION NUMBER: 60/206,848
 ; PRIOR FILING DATE: 2000-05-23
 ; PRIOR APPLICATION NUMBER: 60/207,727
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: 60/230,335
 ; PRIOR FILING DATE: 2000-09-06
 ; PRIOR APPLICATION NUMBER: 60/230,347
 ; PRIOR FILING DATE: 2000-09-09
 ; PRIOR APPLICATION NUMBER: 60/242,578
 ; PRIOR FILING DATE: 2000-10-23
 ; PRIOR APPLICATION NUMBER: 60/253,625
 ; PRIOR FILING DATE: 2000-11-27
 ; PRIOR APPLICATION NUMBER: 60/257,931
 ; PRIOR FILING DATE: 2000-12-22
 ; PRIOR APPLICATION NUMBER: 60/267,636
 ; PRIOR FILING DATE: 2001-02-09
 ; PRIOR APPLICATION NUMBER: 60/269,308
 ; PRIOR FILING DATE: 2001-02-16
 ; Remaining Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 78614
 ; SOFTWARE: Patencin version 3.1
 ; SEQ ID NO 29826
 ; TYPE: DNA
 ; LENGTH: 492
 ; ORGANISM: Neisseria meningitidis
 ; US-10-282-122A-29826

Query Match 23.7%; Score 133.8; DB 13; Length 492;
 Best Local Similarity 60.2%; Pred. No. 7,7e-35;
 Matches 222; Conservative 0; Mismatches 147; Indels 0; Gaps 0;

Qy 49 ATGAACACACCCCTCTGTTTAAAGTGAAGCGGAGTGTCTAACCCGATTGCC 108
 Db 1 ATGCGACATATCTTATCTGTTCTGATCGAAACGATAGGTGCGATGAGCGGTGCTC 60
 Qy 109 GGACTATTGTCGGCGCGGTGTTTAAATGAGACTTGGCGGTGGGTGGCGGAAAG 168
 Db 61 GGTGTTGTTCTTCAACGCGCATTAATATGATTTCTTGGCGGTGGCGGACCGAAGAC 120
 Qy 169 GGGGACGTTTCCGATCAGATGAGTGGTGGCGGGGATGAGAACCATCGAACACTG 228
 Db 121 AAAACCTCTGCTCCGATGACATCTGTACACGCGGACGAGACGTTATCGAATATC 180

Qy 229 ACCAAGCACTGTACAAAGTGTGTTAAGTAAATTAAGTACAGACATCACCGAACTCCC 288
 Db 181 ACCAAGCACTGTAAATTAATTAAGTGTGTTAAGTGTGATTTAAAGGATTTGAGAAAGCGGT 240
 Qy 289 TGTGTGAAAGGGAATGATGCTGTGAAAGGTGAGCGCCATGCCCTTAACCGACGGA 348
 Db 241 TTGTGCAACGGCACTGATGTTGTAAAGTCCGTGCGCGCGGCAAGACCGACGAA 300
 Qy 349 GTGATGAGCTAGCCAGGATTTCCGGGCGCCGATTTGTGATATCTCCGAAACACCCGC 408
 Db 301 TTTTACGCTTACCGCAATCTACCGGGCAGACATCATGACGTAAACGACCGAGCTAT 360
 Qy 409 ACCATCGAA 417
 Db 361 ACGATTGAA 369

RESULT 8

US-09-732-618-5
 ; Sequence 5, Application US/09732618
 ; Patent No. US20010044939A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Abell, Lynn
 ; APPLICANT: Ramodu, Omolayo O.
 ; APPLICANT: Farnodu, Small
 ; TITLE OF INVENTION: Small Subunit of Plant Acetolactate Synthase
 ; FILE REFERENCE: B31435 US NA
 ; CURRENT FILING DATE: 2000-12-08
 ; PRIOR APPLICATION NUMBER: 60/174,437
 ; PRIOR FILING DATE: 2000-01-04
 ; NUMBER OF SEQ ID NOS: 43
 ; SOFTWARE: Microsoft Office 97
 ; SEQ ID NO 5
 ; LENGTH: 1297
 ; TYPE: DNA
 ; ORGANISM: Zea mays
 ; US-09-732-618-5

Query Match 23.5%; Score 132.8; DB 9; Length 1297;
 Best Local Similarity 57.2%; Pred. No. 2,7e-34;
 Matches 279; Conservative 0; Mismatches 207; Indels 2; Gaps 2;

Qy 55 CACACCCCTCTGTTTAAAGTGAAGCGGAGTGTCTAACCCGATTGCGGACCTA 114
 Db 542 CATACCCCTCTCAATCCTTGTAATGATGTGCTGAGTGTCTCAACTTGTAAACAGAGATC 601
 Qy 115 TTGCGCGCGGTGTTTAAATGAGCTTGGCGGTGGGTGGCGGAAACAGGAGGAC 174
 Db 602 TTGCTGCGAGGGGTAAATATAGAGAGGCTTGCTGTTGGCTGAGCTGAGAAAGG 661
 Qy 175 GTTTCGGCATCACATGTTGTTGCGGGGAGTGAAGAACCATGCAACATGACCAAG 234
 Db 662 ATTTCAGTATTAACAAGTGTGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 721
 Qy 235 CAACCTACAGTGTGTTAAGTAAATTAAGTACAGACATCACCGAAATCCCTGTGTC 294
 Db 722 CAGCTTACAGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 781
 Qy 295 GAAAGGAATGATGCTGTGTAAGTGAAGTGAAGCCCAATGCCCTTAACCGAGGAAAGTAT 354
 Db 782 GAAAGGAGCTGATGCTTATTAAGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 841
 Qy 355 GAGCTAGCCAGGATTTCCGGGCGCGCATTTGTGATATCTCCGAGACACCTCA -CCAT 413
 Db 842 GATATGCTGAATTTCCGAGCAAAACCATTTGACGTTTCTTACCATACGTAACCTT 901
 Qy 414 CGAATGTGGGGAACCGGTAATATGTAAGTGAATCTCCAGATGTT -GGCCAATTTGGC 472
 Db 902 CAGCTTACAGATCTTTCAGAGATGTTGCTACTCAAGGTTATTTAGAGGCTATGAG 961
 Qy 473 ATTAAAGGTGCTGTAAGCGGCAAAATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 532

Db 962 ATCTGGAGGTCGCCAGACTGACGAGTGGCACTGGTCGTGAATCGAAGTGCATCC 1021
Qy 533 GAATATCT 540
Db 1022 AAGTACT 1029

RESULT 9

US-10-282-122A-29150
; Sequence 29150, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haeselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyckind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282.122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO: 29150
; LENGTH: 489
; TYPE: DNA
; ORGANISM: Neisseria gonorrhoeae
US-10-282-122A-29150

Query Match 23.4%; Score 132.2; DB 13; Length 489;
Best Local Similarity 59.9%; Pred. No. 2,7e-34;
Matches 221; Conservative 0; Mismatches 148; Indels 0; Gaps 0;
Qy 49 ATGAACACACCCCTCTGTTTGTGAAGATGAGCCGAGGCTTAACCCGCAATGGCC 108
Db 1 ATGGACATATCTTATCTGTTCTGATGAAAAGAAACGATGCGATGAGCCCGCGTGC 60
Qy 109 GACTATTTGGCCCGCGGTGTTTAACTTGAAGCTTGGCGGTGGGTGGCGGAACG 168
Db 61 GGTGTGCTCTGACACCGCATTAACATATGATTTCTTGGCGGTGGCGCGAAGAC 120
Qy 169 GGGGAGCTTCCCGCATCACGATGAGTGGTGGCGGGGATGAGAACACGATCGAACATG 228
Db 121 AAAACCTGTGACGAGTACCATGTTTACCAAGCGGAGGATGATCGAACAATC 180

Qy 229 ACCAAGCACTCTACAACTGTTAGTAAATTAAGTACAGACATCACCGAACTGCC 288
Db 181 ACCAAGCACTCTACAAATTAATTAAGTAAAGTGTGATTTAAAGAAAGCGGT 240
Qy 289 TGTGTGAAAAGGAATTAATGCTGTGTAAGTGAAGCGCCAAATGCCCTAACCGAGCGAA 348
Db 241 TTGTGTGAAAGCGAATGATGTTGTGTAAGATCCCTGCGCCGGAAGACCGGACGAA 300
Qy 349 GTGATGAGCTAGCCGAGATTTCCGCGCCGCAATTTGATCTCGGAAGACCGCTC 408
Db 301 TTTTACGCTTACCGAAATCTACCGGGGACAGCATCATGATGATCCGACCGGAGTAT 360
Qy 409 ACCATCGAA 417
Db 361 ACCATTCGA 369

RESULT 10

US-10-156-761-2719
; Sequence 2719, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIRA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156.761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO: 2719
; LENGTH: 525
; TYPE: DNA
; ORGANISM: Streptomyces avermitilis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(525)
US-10-156-761-2719

Query Match 22.9%; Score 129.6; DB 15; Length 525;
Best Local Similarity 58.2%; Pred. No. 2.2e-33;
Matches 284; Conservative 0; Mismatches 199; Indels 5; Gaps 3;
Qy 52 AAACACACCTCTCTGTTTGTGAAGTGAAGCCGAGTCTTAACCCGCAATGGCGGA 111
Db 7 AAGCACACCTCTCTGCTCTGTCGAGAAACAAGCCGGTGTCTCGCCGGATCACCGCC 66
Qy 112 CTATTTGGCCCGCGGTGTTTAACTTGAAGCTTGGCGGTGGGTGGCGGAACAAGGG 171
Db 67 CTGTTCTCCGCGCGGGGTTTAACTTGAAGCTTGGCGGTGGGTGGCGGAACAAGGG 126
Qy 172 GACGTTTCCCGCATCACCATGATGATGATGATGATGATGATGATGATGATGATGATG 228
Db 127 GACATCTCCCGCATCACCATGATGATGATGATGATGATGATGATGATGATGATGATG 186
Qy 229 ACCAAGCACTCTTACAACTGTTTAACTTGAAGCTTGGCGGTGGGTGGCGGAACAAG 288
Db 187 ACCAAGCACTCTTACAACTGTTTAACTTGAAGCTTGGCGGTGGGTGGCGGAACAAG 246
Qy 289 TGTGTGAAAAGGAATTAATGCTGTGTAAGTGAAGCGCCAAATGCCCTAACCGAGCGAA 348
Db 247 GCGGTGACGCGCAATCTGTTCTGTGTAAGTGGCGCCGCAACGAGACGCGCTCCAG 306
Qy 349 GTGATGAGCTAGCCGAGTATTCGCGCGCGCATTTGTGATATCTCGAAGACACGCTC 408

Db 307 ATCGTCAGATGCTCCAGCTGTTCCGCGCAAGACCGTGGACGTCTCCCGGAGCGCTC 366
 QY 409 ACCATCGA-ATGCTGGGGGACCCGGGTAAATGTAGCAATCTCCAGATGTTGGCCAG 467
 Db 367 ACCATCGAGGCCACCGCGCGGCGACGACAAAGCTGGAGGCCATGCTCAAGATGCTGGAGCG 426
 QY 468 TT-GGCATTAAAGAGGTGGCTCGAAGCGGCAAAATTTGTTTGGTGGGGAATCCGGCGTC 526
 Db 427 TTGGGATCAAGAGAGCTGCTCCAGTCGGAACGATCGGATCGACGTGGGCGACGTTTCG 486
 QY 527 AATAACGA 534
 Db 487 ATCACGGA 494

RESULT 11
 US-10-156-761-1/c
 ; Sequence 1, Application US/10156761
 ; Publication No. US20030119018A1
 ; GENERAL INFORMATION:
 ; APPLICANT: OMURA, SATOSHI
 ; APPLICANT: IKEDA, HARUO
 ; APPLICANT: ISHIKAWA, JUN
 ; APPLICANT: HORIKAWA, HIROSHI
 ; APPLICANT: SHIBA, TADAYOSHI
 ; APPLICANT: SAKAKI, YOSHIYUKI
 ; APPLICANT: HATTORI, MASAHIRA
 ; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
 ; FILE REFERENCE: 249-262
 ; CURRENT APPLICATION NUMBER: US/10/156,761
 ; CURRENT FILING DATE: 2002-05-29
 ; PRIOR APPLICATION NUMBER: JP 2001-204089
 ; PRIOR FILING DATE: 2001-05-30
 ; PRIOR APPLICATION NUMBER: JP 2001-272697
 ; PRIOR FILING DATE: 2001-06-02
 ; NUMBER OF SEQ ID NOS: 15109
 ; SEQ ID NO 1
 ; LENGTH: 9025608
 ; TYPE: DNA
 ; ORGANISM: Streptomyces avermiltilis
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (4187715)
 ; OTHER INFORMATION: a, t, c, g, other or unknown
 US-10-156-761-1

Query Match 22.9%; Score 129.6; DB 15; Length 9025608;
 Best Local Similarity 58.2%; Pred. No. 2.8e-31;
 Matches 284; Conservative 0; Mismatches 199; Indels 5; Gaps 3;
 QY 52 AAACACACCTCTCTGTTTACTGTAAGTGAAGCCGAGTGTCTAACCCGCAATTCGCCGA 111
 Db 3354407 AAGCACACCTCTCTGTTTACTGTAAGTGAAGCCGAGTGTCTAACCCGCAATTCGCCGA 3354348
 QY 112 CTAATTGCCCCCGGCGGTTTAACTGAGAGCTTGGGGGTGGGGGTCCGCGGAACAGGG 171
 Db 3354347 CTGTTCTCCCGGCGGCGGCTTCAACATCGACTGCGCGGCGGTGTCACGAGACCC 3354288
 QY 172 GACGTTTCCCGCATCAACATGTTGTTGCGCGG---GATGAGAACACCATCGAACAACTG 228
 Db 3354287 GACATCTCCCGCATCAACATGTTGTTGCGCGG---GATGAGAACACCATCGAACAACTG 3354228
 QY 229 ACCAAGCACTCTACAACTTGGTTAACTTAAGTACAGGACATCAACGAACTCCC 288
 Db 3354227 ACCAAGCACTCTACAACTTGGTTAACTTAAGTACAGGACATCAACGAACTCCC 3354168
 QY 289 TGTGTGAAAGGAAATGATGCTGTTGAAGTGAAGCGCCAAATGCCCTTAACGAGCGGAA 348
 Db 3354167 GCGGTACAGCGGAGACTCTTTGTGTGAAGTGGCGCGCGCAAGAGACGCGCTCCAG 3354108
 QY 349 GTGATTGAGCTAGCCAGGATATTCGGGCGCGCATTTGTGATATCTTCGGAAGACACCGTTC 408
 Db 3354107 ATGCTGAGATGCTCAGCTGTTCCGCGCAAGACCGTGGAGCTGTCCCGGAGCGCTTC 3354048

QY 409 ACCATCGA-ATGCTGGGGGACCCGGGTAAATGTAGCAATCTCCAGATGTTGGCCAG 467
 Db 3354047 ACCATCGAGGCCACCGCGCGGCGACGACAAAGCTGGAGGCCATGCTCAAGATGCTGGAGCG 3353988
 QY 468 TT-GGCATTAAAGAGGTGGCTCGAAGCGGCAAAATTTGTTTGGTGGGGAATCCGGCGTC 526
 Db 3353987 TTGGGATCAAGAGAGCTGCTCCAGTCGGAACGATCGGATCGGACGTGGGCGACGTTTCG 3353928
 QY 527 AATAACGA 534
 Db 3353927 ATCACGGA 3353920

RESULT 12
 US-09-732-618-15
 ; Sequence 15, Application US/09732618
 ; Patent No. US2001004939A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Abell, Lynn
 ; APPLICANT: Falco, Carl
 ; APPLICANT: Famodu, Omolayo O.
 ; TITLE OF INVENTION: Small Subunit of Plant Acetolactate Synthase
 ; FILE REFERENCE: B1435 US NA
 ; CURRENT APPLICATION NUMBER: US/09/732,618
 ; CURRENT FILING DATE: 2000-12-08
 ; PRIOR APPLICATION NUMBER: 60/174,437
 ; PRIOR FILING DATE: 2000-01-04
 ; NUMBER OF SEQ ID NOS: 43
 ; SOFTWARE: Microsoft Office 97
 ; SEQ ID NO 15
 ; LENGTH: 1435
 ; TYPE: DNA
 ; ORGANISM: Oryza sativa
 US-09-732-618-15

Query Match 22.8%; Score 129; DB 9; Length 1435;
 Best Local Similarity 56.6%; Pred. No. 5.7e-33;
 Matches 277; Conservative 0; Mismatches 210; Indels 2; Gaps 2;
 QY 54 AACACCCCTCTCTGTTTACTGTAAGTGAAGCCGAGTGTCTAACCCGCAATTCGCCGACT 113
 Db 629 AATACCTCTATCCATCTTGTCAATGATGCTCTGTTGTTCTCAACATTTGTAAGGGGT 688
 QY 114 ATTTGCCCCCGGCGGTTTAACTGAGAGCTTGGGGGTGGGGGTCCGCGGAACAGGGGGA 173
 Db 689 CTTTGCTCGAGAGGCTACATATPACAGAGTCTTGGCTGTAAGGCCAGCTGAAGAGTCAGG 748
 QY 174 CGTTTCCCGCATCAACATGTTGTTGCGGGGATGAGAACCATCGAACACTGACCA 233
 Db 749 CTTTCCCGTATTAACAACATGTTCTCTGGAACAGATGATTCATTGAGAGTTAATTCA 808
 QY 234 GCAACTTACAACTGTTGTTAACTTAAGTAAAGTACAGACATCAACGAACTCCGTGTG 293
 Db 809 GCAGCTTACAACTGTTGTTATGATGATGATGATGATGATGATGATGATGATGATGATG 868
 QY 294 GGAAGGAAATGATGCTGTGTAAGGTGAGGCGCAATGCCCTTAACCGGAGCGGAAGTAT 353
 Db 869 TGAAGAGAACTTATGCTTATCAAGGTTCTGTGAACACTGCTGCTGAGAGACATTA 928
 QY 354 TGAGCTAGCCAGGATATTCGGGCGCGCATTTGATGATGATGATGATGATGATGATGATG 413
 Db 929 AGATATGCTGAATCTTCCGGGCAAAATCTGTTGATGTTCTGATCAACTGTTACGTT 988
 QY 414 -CGAATGTTGGGACCCCGGTTAAATGTTAGCAATCTTCAGATGTTGG-CCAACTGG 471
 Db 989 ACAGCTTACGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1048
 QY 472 CATTAAGAGGTGCTCGAAGCGGCAAAATTTGTTTGGTGGGGAATCCGCGCTCAATAC 531
 Db 1049 CATGTGAGGTGCGCAAGACAGGCGAGTGGCGCTGTTGCGGAATCCGTTGATTC 1108
 QY 532 GGAATATCT 540


```

Db      61 GGTCTTTTTCGGACCGCGCTACACATTGAAACCTTAGCGTGGCTCCGACCGAAGAC 120
Qy      169 GGGAGCGTTTCCCGCATCAACATGTGTGTCGGGGGATGAGAACACCATCGAACACTG 228
Db      121 CATTGCTGTCCGCGATGACCATCGTCTCCATTGGGTGGAGAGAGTATCGAACAGATC 180
Qy      229 ACCAGCACTCTACAGTTGGTTAACTTAAGTACAGACATCACCGAACTCC 288
Db      181 ACCAAGCATCTTAACCGCTGATCGAGTGTGAAAGTGTGACCTTAACCGAGGCGCC 240
Qy      289 TGTGTGAAAGGAAATTGATGCTGTGTAAGTGAAGCCAAATGCCCTTAACCGAGCGAA 348
Db      241 CACATGAGCGGAGGTGATGTTGATCAAGTGAAGGCGGTGGCAAGGACGTAGAG 300
Qy      349 GTGATTGACCTAGCCAGGTATTCGGGGCCCGCATTTGGATATCTCGAAGACCGCTC 408
Db      301 ATGAAGCGATGTGCGATATTTTCGGCGCGCCCATCATGACGTACCGAAAGACTTAC 360
Qy      409 ACCATGAAATGGTGGGGGACCGCGGTAAA 437
Db      361 ACATGAACTGACGGGCGGACGACACA 389

```

RESULT 15

```

US-10-282-122A-11953
; Sequence 11953, Application US/10282122A
; Publication No. US20040029129A1

```

GENERAL INFORMATION:

```

; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Cair, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.

```

```

; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

```

```

; FILE REFERENCE: ELITRA.034A

```

```

; CURRENT APPLICATION NUMBER: US/10/282,122A

```

```

; PRIOR FILING DATE: 2003-02-20

```

```

; PRIOR APPLICATION NUMBER: 60/191,078

```

```

; PRIOR FILING DATE: 2000-03-21

```

```

; PRIOR APPLICATION NUMBER: 60/206,848

```

```

; PRIOR FILING DATE: 2000-05-23

```

```

; PRIOR APPLICATION NUMBER: 60/207,727

```

```

; PRIOR FILING DATE: 2000-05-26

```

```

; PRIOR APPLICATION NUMBER: 60/230,335

```

```

; PRIOR FILING DATE: 2000-09-06

```

```

; PRIOR APPLICATION NUMBER: 60/230,347

```

```

; PRIOR FILING DATE: 2000-09-09

```

```

; PRIOR APPLICATION NUMBER: 60/242,578

```

```

; PRIOR FILING DATE: 2000-10-23

```

```

; PRIOR APPLICATION NUMBER: 60/253,625

```

```

; PRIOR FILING DATE: 2000-11-27

```

```

; PRIOR APPLICATION NUMBER: 60/257,931

```

```

; PRIOR FILING DATE: 2000-12-22

```

```

; PRIOR APPLICATION NUMBER: 60/267,636

```

```

; PRIOR FILING DATE: 2001-02-09

```

```

; PRIOR APPLICATION NUMBER: 60/269,308

```

```

; PRIOR FILING DATE: 2001-02-16

```

```

; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 11953
; LENGTH: 525
; TYPE: DNA
; ORGANISM: Burkholderia cepacia
; US-10-282-122A-11953

```

```

Query Match      22.3%; Score 126; DB 13; Length 525;
Best Local Similarity 57.5%; Pred. No. 3,7e-32;
Matches 245; Conservative 0; Mismatches 180; Indels 1; Gaps 1;

```

Qy	48	CATGAACACACCTCTCTCTTTTAACTTAAGTGAAGCCGAGTGTAACTTCCGATTC	107
Db	36	CATGAACACATATTTTCTCTCTCTGAGAACGAAACCGGCGCGCTGTGTGCGCTGT	95
Qy	108	CGGACTATTTTCCCGCGGTGTTTAACTTGAAGTGTGCGGTGGGTGCGCGGACA	167
Db	96	CGGCTTTTTCGCGACGCGGCTAACACATCGAAACCTTGAAGTGTGCGCGGACG	155
Qy	168	GGGGAGCTTTCCCGATCAACATGTTGTGTCGGGGGATGAGAACACATCGAACACT	227
Db	156	TCAATCGCTGTGCGGCTCACATCGTTCCATTGGCTCCGACGAGTGTATCGAACAGAT	215
Qy	228	GACCAAGCACTCTACAAATTTGTTAACTTAAGTACAGACATCACCGAACTCC	287
Db	216	CACAGACATCTGAACCGCTTATCGAGTGTGAAGTGTGAACCTGACGACGCTC	275
Qy	288	CTGTGTGAAAGGAAATTGATGCTGTGAAGTGAAGGCCAATGCCCTTAACGAGCGGA	347
Db	276	ACACATGGAACGCGAGCTGATGCTGATCAAGTACGTCAGTGGGCAAGGAGCGGAGA	335
Qy	348	AGTGAATGACCTAGCCAGGTATTCGGGGCCGCTTGTGATATCTCGAAGACACCGT	407
Db	336	AATGAAGCGATGTGGAATTTTCGGGCGCATCATGACGTACCGAAGACCTTA	395
Qy	408	CACCATGAATGTGTGGGAGACCGG-GGTAAATGTAGCATCTCCAGATGTTGGCAA	466
Db	396	CACGATGAATTTGACGGGCGCGAGCAAACTCGACGATTCATCGAGGGGCTGAGACG	455
Qy	467	GTGGC 472	
Db	456	GGGCGC 461	

Search completed: July 25, 2004, 13:28:33
 Job time : 281.203 secs

Mon Jul 26 12:13:49 2004

us-09-893-033-17.rst

Page 1

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 25, 2004, 03:59:07 ; Search time 1611.95 Seconds
(without alignments)

10466.900 Million cell updates/sec

Title: US-09-893-033-17
Perfect score: 565
Sequence: 1 gtggaatttaccacatg.....ccctgaatccaagttag 565

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 segs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
EST:*
1: em_escba:*
2: em_escin:*
3: em_escin:*
4: em_escin:*
5: em_escov:*
6: em_escpl:*
7: em_escro:*
8: em_escro:*
9: gb_esc1:*
10: gb_esc2:*
11: gb_esc3:*
12: gb_esc3:*
13: gb_esc4:*
14: gb_esc5:*
15: em_escfun:*
16: em_escfun:*
17: em_escfun:*
18: em_escfun:*
19: em_escfun:*
20: em_escfun:*
21: em_escfun:*
22: em_escfun:*
23: em_escfun:*
24: em_escfun:*
25: em_escfun:*
26: em_escfun:*
27: em_escfun:*
28: gb_esc1:*
29: gb_esc2:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	142	25.1	679	13	BQ407261 GA_Ed010
2	137	24.2	782	14	CB642034 OSJNEB01P
3	136	24.1	698	14	CB647745 OSJNEB01M
4	131.6	23.3	1380	11	AY105043 Zea mays

C	5	130.6	23.1	911	14	CA238460	SCRPL503
C	6	129.8	23.0	799	14	CK291478	EST754192
C	7	129.8	23.0	962	14	CK297674	EST760388
C	8	129.4	22.9	867	14	CK243941	EST727578
C	9	129	22.8	584	14	CA242014	SCRPL308
C	10	129	22.8	626	14	CB672821	OSUNE07B
C	11	129	22.8	695	14	BP184863	BP184863
C	12	128.8	22.8	697	14	CF210202	CAB20006
C	13	128.6	22.8	548	14	CD429148	CD429148
C	14	128.6	22.8	782	14	BG581436	EST483170
C	15	128.2	22.7	891	14	CK243942	CK243942
C	16	128.2	22.7	938	14	CK257587	EST741234
C	17	128.2	22.7	972	14	CK243965	CK243965
C	18	128.2	22.7	977	14	CK255153	EST738790
C	19	127.8	22.6	664	12	BJ574968	BJ574968
C	20	127.8	22.6	711	12	BJ574939	BJ574939
C	21	127.6	22.6	541	10	BE999544	EST41267
C	22	127.4	22.5	728	14	CB978073	CB978073
C	23	126.6	22.4	885	14	CK258640	CK258640
C	24	126.4	22.4	638	10	BE823545	BE823545
C	25	125.8	22.3	684	14	CB654393	CB654393
C	26	125.6	22.2	710	14	CF397385	RTD3 3 B
C	27	124.6	22.1	703	14	CF480319	POL1 65 B
C	28	124.2	22.0	671	13	BU102997	SCPECL602
C	29	124.2	22.0	671	13	CA097418	CA097418
C	30	124.2	22.0	676	13	BU102998	BU102998
C	31	124.2	22.0	676	13	CA097419	CA097419
C	32	123.4	21.8	660	12	BP184425	BP184425
C	33	123.4	21.8	824	14	CB654392	CB654392
C	34	123.4	21.7	601	10	AM696137	AM696137
C	35	122.6	21.7	589	14	AM181183	687021E04
C	36	122.2	21.6	589	14	CD909456	CA68 112K
C	37	121.6	21.5	632	13	BQ415906	BQ415906
C	38	121	21.4	598	14	CD830820	BN40 0.460
C	39	121	21.4	811	14	CK194674	FGAS00310
C	40	120.4	21.3	552	12	BJ310586	BJ310586
C	41	119.8	21.2	606	13	CA078461	SCRLAM100
C	42	119.2	21.1	797	12	BG836929	2M08 0940
C	43	118.8	21.0	702	12	BU169282	BU169282
C	44	118.8	21.0	702	12	BU581766	BU581766
C	45	118.6	21.0	563	13	CA021505	H240B18r

ALIGNMENTS

RESULT 1
BQ407261 679 bp mRNA linear EST 22-May-2002
LOCUS GA_Ed0104F04f Gossypium arboreum 7-10 dpa fiber library Gossypium
DEFINITION GA_Ed0104F04f Gossypium arboreum 7-10 dpa fiber library Gossypium
arboresum cDNA clone GA_Ed0104F04f, mRNA sequence.

ACCESSION BQ407261 GI:21094948
VERSION BQ407261.1 GI:21094948

KEYWORDS
SOURCE
ORGANISM

Gossypium arboreum
Gossypium arboreum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Malvales; Malvaceae; Malvoideae; Gossypium.

REFERENCE
Wing, R.A., Fritsch, D., Yu, Y., Main, D., Rambo, T., Simmons, J.,
Henry, D., Wood, T.C., Leslie, A. and Wilkins, T.A.
An integrated analysis of the genetics, development, and evolution
of the cotton fiber
Unpublished (2000)
JOURNAL
COMMENT
Contact: Wing RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: rwing@clemson.edu
Total High Quality bases = 441

Seq primer: TAAACGACTCATTATAGCG
High quality sequence start: 6
High quality sequence stop: 620.
Location/Qualifiers

FEATURES

1..679
/organism="Gossypium arboreum"
/mol_type="mRNA"
/strain="AKA"
/cultivar="8400"
/db_xref="taxon:29729"
/clone="GA_Ed0104F04f"
/tissue_type="Fibers isolated from bolls harvested 7-10 dpa"
/lab_host="E. coli"
/clone_lib="Gossypium arboreum 7-10 dpa fiber library"
/note="Vector: pBK-CMV; Site_1: EcoRI; Site_2: XhoI"

ORIGIN

Query Match 25.1%; Score 142; DB 13; Length 679;
Best Local Similarity 57.0%; Pred. No. 2.5e-29;
Matches 259; Conservative 0; Mismatches 195; Indels 0; Gaps 0;

QY 55 CACACCTCTCTGTTTAAAGTGAAGATGAGCCGAGTGTAAACCCGATTCGGACTA 114
DB CACTTATTCATCTGTGTAATGACTCTCCGGAGTTCTAAACCTGTACAGGTGT 285
QY 115 TTTCGCCGCCGCTTTTAAACATTGAGCTTGGCGGTGGGCTGGGCGAACAAGGGGAC 174
DB TTTCGTCGAAGGGCTAATATCTTCAGATTGGCTGTGGAACATGCTGAAGTTAGGGG 345
QY 175 GTTTCGCCGATCACCATGATGATGCGGGGAGTGAAGAACACCATCGAACAATGACCAAG 234
DB CTTTCCGATTAACAACATGTTGTCTCTGTGTAAGTGAATCAATTAGCAATTTAGTCAG 405
QY 235 CACTCTTACAACTTGTATTAAGTATTAAGTACAGGACATCACCAGAACTCCCTGTGTG 294
DB CAACCTTAATAGCTGTGATGATATGATGATGATGATGATGATGATGATGATGATGAT 465
QY 295 GAAAGGAAATGATCTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 354
DB GAACGAGAACTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 525
QY 355 GAGCTAGCCCAAGATTTCCGGGCCCGCATTTGTGATATCTCCGAAGACACCGTACCATC 414
DB GACATTTGCCAATTTTAAAGGCGCAAGCTGTGATGATGATGATGATGATGATGATGAT 585
QY 415 GAATGATGAGGAGACCCGGGTAAATGATGATGATGATGATGATGATGATGATGATGAT 474
DB GAGCTACAGAGATTTAACAATAATGATGATGATGATGATGATGATGATGATGATGAT 645
QY 475 TAAAGAGGTGCTCGAACGGGCAAAATTTGCTTTG 508
DB CTGTGAAGTTGCTAAACTGGAACCTCGCACTG 679

RESULT 2
CB642034/c 782 bp mRNA linear EST 08-APR-2003

LOCUS CB642034
DEFINITION OSJUNBD1P24.r OSJUNBD Oryza sativa (japonica cultivar-group) cDNA
OSJUNBD1P24.3', mRNA sequence.

ACCESSION CB642034
VERSION CB642034.1 GI:29637025

KEYWORDS EST.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Eubartoideae; Oryzaceae; Oryza.
1 (bases 1 to 782)
Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,
Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.
TITLE Large-scale identification of ESTs involved in the interaction
between rice and Magnaporthe grisea

JOURNAL

Unpublished (2003)
Contact: Rod Wing
Arizona Genomics Institute
University of Arizona
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
85721-0088, USA
Tel: 520 626 3967
Fax: 520 621 9288
Email: http://genome.arizona.edu

FEATURES

Seq primer: gga aac agc tat gac cat g.
Location/Qualifiers

ORIGIN

1..782
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="OSJUNBD1P24"
/tissue_type="leaf"
/dev_stage="3 week"
/lab_host="DH10B"
/clone_lib="OSJUNBD"
/note="Vector: pBluescript II KS +; Site_1: EcoRI; Site_2: XhoI; 24 hrs after inoculation with Rice Blast (Che 86061)"

Query Match 24.2%; Score 137; DB 14; Length 782;
Best Local Similarity 57.7%; Pred. No. 7.1e-28;
Matches 282; Conservative 0; Mismatches 205; Indels 2; Gaps 2;

QY 54 AACACCCCTCTCTGTTTAAAGTGAAGATGAGCCGAGTGTAAACCCGATTCGGACTA 113
DB AACTTCTTGTCAATTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 640
QY 114 ATTTCGCCGCCGCTTTTAAACATTGAGAGCTTGGCGGTGGGCTGGGCGAACAAGGGGA 173
DB TTTCGCCGAGAGGCTCAATATTCAGAGTCTGCTGTGGCCAGCTGAAAAAGAAAG 580
QY 174 CTTTCCGATCACCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 233
DB CACTTCTCGATCACTACTGTTGTCTCCGGAACGTATGATGATGATGATGATGATGAT 520
QY 234 GCAACTCTCAAGTGTGTTAAGTATTAAGTACAGACATCAACCGAACTCCCTGTGT 293
DB CCACTGTAATAGCTCATTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 460
QY 294 GGAAGGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 353
DB TCTTAGAGAGTAAAGATATTAAGATTTGCTGTTAAACACACAGCCCGCAGGCTATCT 400
QY 354 TGAGCTAGCCGAGTATTCGGGCCCGCATTTGTGATATCTCCGAAGACACCGTAC-CA 412
DB AGATTTGCTGATATTTTCCGGCCCAAACTGTGATGATGATGATGATGATGATGATGAT 340
QY 413 TCGAATGATGAGGAGACCCGGGTAAATGATGATGATGATGATGATGATGATGATGATGAT 471
DB TCAGCTTACTGAGACCTTGATTAATGATGATGATGATGATGATGATGATGATGATGAT 280
QY 472 CATTAAGAGGTGCTCGAAGCGGCAAAATTTGCTTTGATGATGATGATGATGATGATGAT 531
DB CATTTGTGAGTTGACGAACTGGGAGGATGCTGTTGCGCGGTGATGATGATGATGAT 220
QY 532 GGAATATCT 540
DB 219 CAATATCT 211

RESULT 3
CB647745/c

LOCUS CB647745 698 bp mRNA linear EST 08-APR-2003
 DEFINITION OSUNEb10M2.1 r OSNEb Oryza sativa (japonica cultivar-group) cDNA
 ACCESSION CB647745
 VERSION CB647745.1 GI:29642738
 KEYWORDS EST.
 SOURCE Oryza sativa (japonica cultivar-group)
 ORGANISM Oryza sativa (japonica cultivar-group)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Ehrhartoideae; Oryzoideae; Oryza.
 1 (bases 1 to 698)
 Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,
 Kudrna,D., Dean,R., Soderlund,C., Wang,R. and Wang,G.
 Large-scale identification of ESTs involved in the interaction
 between rice and Magnaporthe grisea
 Unpublished (2003)
 CONTACT: Rod Wang
 Arizona Genomics Institute
 University of Arizona
 Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
 85721-0088, USA
 Tel: 520 626 3967
 Fax: 520 621 9288
 Email: http://genome.arizona.edu
 PCR Primers
 FORWARD: gta aac cga cgg cca gtc
 BACKWARD: gga aac agc tat gac cat g
 Plate: 10 row: M column: 22
 Seq primer: gga aac agc tat gac cat g.
 Location/Qualifiers
 1..698
 /organism="Oryza sativa (japonica cultivar-group)"
 /mol_type="mRNA"
 /cultivar="Nipponbare"
 /db_xref="taxon:39947"
 /clone="OSUNEb10M2"
 /issue_type="leaf"
 /dev_stage="3 week"
 /lab_host="DH10B"
 /clone_lib="OSUNEb"
 /note="Vector: pBluescript II KS +, Site 1: EcoRI, Site 2:
 XhoI; 24 hrs after inoculation with Rice Blast (Cne
 86061)"

ORIGIN
 Query Match 24.1%; Score 136; DB 14; Length 698;
 Best Local Similarity 57.6%; Pred. No. 1.3e-27;
 Matches 281; Conservative 0; Mismatches 205; Indels 2; Gaps 2;

55 GACACCCCTCTCTGTTTAAAGTGAAGCGGAGTCTAACCCGATTGCCGACTA 114
 |||
 698 CATCTCTTCCATTTGTGATGATTTCCCTGAGTCTCAATGTTGMAACAGTGT 639
 |||
 115 TTGCGCGCGGTGTTTAACTAGAGAGCTTGGCGGTGGCGGACAGGGGAC 174
 |||
 638 TTCTCCCGAAGGCTACATATTCAGAGTCTGGCTGTGGCCAGCTGAAAATAAGGC 579
 |||
 175 GTTTCGCCGATCACCATGTTGTGCGGGGAGTGAACACCCATGAACTGACAC 234
 |||
 578 ACTTCTCGGATCAGTACTGTGTCCCTGGAAGTGAAGTATTTGCAAGTATTAAC 519
 |||
 235 CAACCTTACAGTTGGTTAACTAATTAAGTACGAGCATACCGAATCTCCCTGTG 294
 |||
 518 CAACGTATTAACCTAATGATTTTATGAGTCCGAGATCTTACCATTTTACATTCT 459
 |||
 295 GAAAGGAAATGATGTTGAGTGAAGTGAAGCGCCAAATGCCCTAACCGAGCGAATGATT 354
 |||
 458 GCTAGAGATTAATGATCATTAAGATGCTGTAAACACAGCCCGAGAGCTTCTCTA 399
 |||
 355 GAGCTAGCCAGGATTTCCGGGCGCGCATTTGGATATCTCCGAAGACCGCTAC-CAT 413
 |||
 398 GATATTGCTGATATTTCCGGGCGCAAACTGTGATGATGATGATGATGATGATGATGAT 339

LOCUS AY105043 1380 bp mRNA linear HTC 16-OCT-2002
 DEFINITION Zea mays PC0069007 mRNA sequence.
 ACCESSION AY105043
 VERSION AY105043.1 GI:21208121
 KEYWORDS HTCC.
 SOURCE Zea mays
 ORGANISM Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoidae; Andropogoneae; Zea.
 1 (bases 1 to 1380)
 Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitlitt,M.S.,
 Arthur,L.W., Hanaley,M., Morgante,M. and Tinsley,S.V.
 Maize Mapping Project/Dupont Consensus Sequences for Design of
 Overgo Probes
 Unpublished (2002)
 2 (bases 1 to 1380)
 Coe,E.H.
 Direct Submission
 Submitted (25-APR-2002) Maize Mapping Project, University of
 Missouri, Columbia, MO 65211, USA
 If you are interested in getting corresponding physical clones,
 these are publicly available from ZMDB and may be found by BLAST
 searching at MSU, maizemap.org; ZMDB, www.zmdb.iastate.edu; TRGR,
 www.trgr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the
 maize cDNA sequences is either Virginia Walbot, Stanford or Pat
 Schnable, Iowa State, then clones may be requested from ZMDB:
 www.zmdb.iastate.edu.
 Location/Qualifiers
 1..1380
 /organism="Zea mays"
 /mol_type="mRNA"
 /db_xref="taxon:4577"
 /db_xref="maizedb:633742"
 /clone_lib="Maize Mapping Project/Dupont Consensus
 Library"
 /note="this sequence is part of a project of EST
 assemblies resulting from the application of public
 contigs to seed Dupont contigs; this resource was
 assembled by Dupont as part of a collaboration for the
 overgo addressing of BACs in conjunction with the Maize
 Mapping Project"

ORIGIN
 Query Match 23.3%; Score 131.6; DB 11; Length 1380;
 Best Local Similarity 57.3%; Pred. No. 3.3e-26;
 Matches 276; Conservative 0; Mismatches 204; Indels 2; Gaps 2;

61 CTCTCTGTTTAAAGTGAAGCGGAGTCTAACCCGATTGCCGACTATTGCC 120
 |||
 635 CTCTCCATCTCTTGAATGACTGTCTGCTGCTCAACTTTTAAACAGAGTCTTGTCT 694
 |||
 121 CGCGGTGTTTAACTGAGAGCTTGGCGGTGGCGGTGGCGGACAGCGGGAGCTTCC 180
 |||
 695 CGCGGGGCTTACATATTAACAGAGCTTGTCTGTGGCCCACTGAAGAGAGGAGATTTCG 754
 |||
 181 CGCATCACCATGTGTGTGCGGGGAGTGAACAACCATCGAACACTGAACCAACTC 240

Db 755 CGGATTACACAGTGTCTCTCTGTAATCTGTAATCCATGAGAAAGTTAGTTCACACCTT 814

Qy 241 TACAAAGTGTAACTGTAATTAAGTACAGACATCACCGAACTCCCTGTGTGAAAG 300

Db 815 TACAAAGTGTAACTGTAATTAAGTACAGACATCACCGAACTCCCTGTGTGAAAG 874

Qy 301 GAATGATGCTGTGTAAGTGTGAGCCCAATGCCCTTACCGAGCGGAAGTATGACTA 360

Db 875 GAATGATGCTGTGTAAGTGTGAGCCCAATGCCCTTACCGAGCGGAAGTATGACTA 934

Qy 361 GCCAAGTATTCGGGCGCCGATGTGTGATATCTCCGAAGACCGGTAC-CATGGAATG 419

Db 935 GGTGAATCTTCGCGCAAAACCTGTGATGTTCTGACACACAGTAACGCTTCAGCTT 994

Qy 420 GTGGGAGACCGCGGTAATGAGTGAATCTCCAGATGTTGG-CCAAGTTGGATTA 478

Db 995 ACTGAAGATCTTGAAGAGATGTTGACCTACAAAGCTTATGGACCTATGGATCTGT 1054

Qy 479 GAGTGAGTGTGAAACGGGCAAAATGCTTGTGTGCGGGAATCCGGCTCAATACGGAATAT 538

Db 1055 GAGTGCGCAAGTGAAGAGATGACCTGTGTCGGCAATCGAAGTGTGACTCCAGTAC 1114

Qy 539 CT 540

Db 1115 CT 1116

RESULT 5
CA238460/c 911 bp mRNA linear EST 25-SEP-2003
LOCUS SCRFPL5036A10.g Saccharum officinarum FUS Saccharum officinarum
DEFINITION cDNA clone SCRFPL5036A10 5', mRNA sequence.
ACCESSION CA238460
VERSION CA238460
KEYWORDS EST.
SOURCE Saccharum officinarum
ORGANISM Saccharum officinarum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Saccharum.
REFERENCE 1 (bases 1 to 911)
Vettore, A.L., da Silva, F.R., Kemper, E.L. and Arruda, P.
The libraries that made SUCEST
Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
COMMENT Contact: Arruda, P.
Centro de Biologia Molecular e Engenharia Genetica
Universidade Estadual de Campinas
Caixa Postal 6010, 13083-970, Campinas SP, Brazil
Tel: 55 19 3788 1137
Fax: 55 19 3788 1089
Email: parnada@unicamp.br
Clone distribution: clone distribution information can be found
through the Brazilian Clone Collection Center (BCCC) at
http://www.bcccenter.fcav.unesp.br
Plate: 036 row: A column: 10
Seq primer: T7 Promoter Primer.
FEATURES
source location/Qualifiers
1..911
/organism="Saccharum officinarum"
/mol_type="mRNA"
/db_xref="taxon:4547"
/clone="SCRFPL5036A10"
/lab_host="DH10B"
/note="Organ: Developed inflorescence (20cm-long) without
rachis; Vector: pSPori1; Site: 1: SalI; Site: 2: NotI; An
unidirectional cDNA library generated from [Developed
inflorescence (20cm-long) without rachis]. cDNA was
prepared from polyA+ mRNA using SuperScript Plasmid
System Kit (Invitrogen). The double-strand cDNAs were
fractionated in a sepharose CL-2B 40cm-columns and
fragments sizing between 0.8 and 1.5 Kb were

directionally cloned into the vector. Details of each
source of RNA and library construction can be obtained at
<http://succest.fad.ic.unicamp.br/public>

ORIGIN

Query Match 23.1%; Score 130.6; DB 14; Length 911;
Best Local Similarity 56.9%; Pred. No. 5, 2e-26;
Matches 275; Conservative 0; Mismatches 206; Indels 2; Gaps 2;

Qy 60 CCTCTCTGTTTAAAGTGAAGTGAAGCCGAGTGTAAACCGCATTTGCCGACTTTTTC 119

Db 646 CCTCTCTCTGTTTAAAGTGAAGTGAAGCCGAGTGTAAACCGCATTTGCCGACTTTTTC 587

Qy 120 CCGCCGTGTTTAAACATTTAGAGCTTGGCGGTGCGGTGCGGGAACGGGAGCGTTTC 179

Db 586 TGGCAGNGGTGCAATATATCAAGAGCTTGTCTTGGCCAGTGAAGGAAGCATTTTC 527

Qy 180 CCGCATCACATGTGTGTGTGCGGGAAGTGAAGAACCATTCGAACACTGACCAACT 239

Db 526 GCGTATTACAAAGTGTGTTCTCGTGAACCTTGAATCCATTGGAAGTTAGTTCAAGACT 467

Qy 240 CTACAAAGTGTGTTAAAGTAAATTAAGTACAGACATCCAGAACTCCCTGTGTGGAAG 299

Db 466 TTACAAAGTGTGTTAGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 407

Qy 300 GGAATGATGCTGTGTAAGTGAAGTGAAGCGGCAATGCCCTTAACCGAGCGAAGTATGACT 359

Db 406 GGAATGATGCTGTGTAAGTGAAGTGAAGCGGCAATGCCCTTAACCGAGCGAAGTATGACT 347

Qy 360 AGCCGAGATTCGCGGCGCGCGCGCATTTGTGATATCTCCGAAGACCGGTAC-CATGGAAT 418

Db 346 TGCGAAATCTTCGCGCAAAACCTGTGTGATTTTCTGACCTTAAGTGAAGTGAAGCTTCACT 287

Qy 419 GGTGGGGAACCGGGAATTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 477

Db 286 TACTGAAGTATTTGAAGAGATGTTGACCTACAAAGTATTTAGGAGCATATGGCATCTG 227

Qy 478 AGAGTGTGCTGAAGCGGCAAAATGCTTGTGTGTGCGGGAATCCGCGCATATCGAAT 537

Db 226 CAGGTGCGCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 167

Qy 538 TCT 540

Db 166 CCT 164

RESULT 6
CK291478 799 bp mRNA linear EST 15-DEC-2003
LOCUS EST754192 Nicotiana benthamiana mixed tissue cDNA library,
DEFINITION normalized, full-length Nicotiana benthamiana cDNA clone NBNC250 5',
end. mRNA sequence.
ACCESSION CK291478
VERSION CK291478
KEYWORDS EST.
SOURCE CK291478.1 GI:39871957
ORGANISM Nicotiana benthamiana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons, core eudicots;
asterids; Lamiales; Solanales; Solanaceae; Nicotiana.
REFERENCE 1 (bases 1 to 799)
Buell, C.R., Hart, A., Ziemann, V., Kazemiycheva, S.A., Day, B.,
Staskawicz, B., Jin, H. and Baker, B.
Generation of EST sequences from Nicotiana benthamiana
Unpublished (2003)
COMMENT Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@tigr.org
Clones can be requested from TIGR via potato@tigr.org
Seq primer: ART TAG GTG ACA CTA TAG.
FEATURES
source location/Qualifiers
1..799

/organism="Nicotiana benthamiana"
/mol_type="mRNA"
/db_xref="taxon:4100"
/clone="NBYC250"
/tissue_type="abiotic and biotic stress-treated leaves,
callus tissue and root tissue"
/lab_host="DH10B-Tona"
/clone_lib="Nicotiana benthamiana mixed tissue cDNA
library, normalized, full-length"
/note="Vector: pCMVSPORT6.1; Site_1: EcoRI; Site_2: NotI;
supplier: RNA was isolated from Nicotiana benthamiana
tissues that include callus, roots from liquid culture
grown plants, heat-stressed leaves (38 C, 3 hr and 6 hr),
cold-stressed leaves (5 C 3 hr, 6hr), and pathogen
challenged leaves (Pseudomonas syringae pv tomato 12 hr;
Xanthomonas campestris pv campestris 12 hr, 18hr;
Pseudomonas syringae pv phaseolicola 18hr, and Xanthomonas
campestris pv vesicatoria 18hr). RNA was isolated from
these tissues and pooled in approximately equal molar
amounts."

ORIGIN

Query Match 23.0%; Score 129.8; DB 14; Length 799;
Best Local Similarity 58.1%; Pred. No. 8,2e-26;
Matches 270; Conservative 0; Mismatches 187; Indels 8; Gaps 2;
QY 55 CACACCCCTCTGTTTGTAGTGAAGTGAAGCCGAGTCTAACCCGATTGCCGACTA 114
DB 280 CACACCATTTTCAGTATTTTGTGGGATGAAGTGTATATTAACAGAAATTGCTGAGTT 339
QY 115 TTTGCCCGCGTGTGTTTAACTTGAAGCTTGGCGGTGGCGGACAGGGGAC 174
DB 340 TTTGCTCGAAGGGGTTAACAATATGATCACTTGTCTGTGGTTGACAA-----GGAT 393
QY 175 GTTTCGCCGATCAGCATGTTGTCGGGGGATGAACACATCGAACACTGACCAAG 234
DB 394 AAGCTCTCTTACCATAGTGTCTCTGGAACAGGAAGTCTTGCAACAAGTTGTAGAG 453
QY 235 CAAGCTCTCAAGTTGTTAAGCTTAATTAAGTACAGACATCACCGAAATCTCCCTGTG 294
DB 454 CAGCTCAACAAGCTTGGAAGTCTTGAAGTGAAGGATCTGTCAAGAGAACCAAGTT 513
QY 225 GAAAGGGAATTGATCTGTGTGAAGTGAAGCCGCAATGCCCTAACCGGCGGAAGTGT 354
DB 514 GAACGAGATTGATCTCATTAAGCTGAATGACATGACAGACAAAGCTGAATTTATG 573
QY 355 GAGTAGCCCAAGTATTCGGGCGCCGCAATTGATATCTCGAAGACACGTCACCATC 414
DB 574 TGGTTGTGACATCTTCAGAGGGAATTTGTGATATATAGACACTTTTGTGACTATC 633
QY 415 GAATGG--TGGGGACCCCGGTAAATGTGTAGCAATCTCCAGATTTGGCCAACTTGGC 472
DB 634 GAGGTGACTGTGACCCGAGGAAATGTGCTGTTTGAAGAAATTTAAGCAAGTTGGG 693
QY 473 ATTAAGAGGTGCTCGAACGGGCAAAATTTGCTTTGGTGGCGGAA 517
DB 694 ATCAAAAGACTAGCCCGAAGCTGAAAGATGCTTTGAGACGTGA 738

RESULT 7
CK297674 962 bp mRNA linear EST 15-DEC-2003
LOCUS ESTT60388 Nicotiana benthamiana mixed tissue cDNA library,
DEFINITION normalized, full-length Nicotiana benthamiana cDNA clone NBMDS4 5'
END, mRNA sequence.

ACCESSION CK297674
VERSION CK297674
KEYWORDS EST.
SOURCE Nicotiana benthamiana
ORGANISM Nicotiana benthamiana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamiales; Solanales; Solanaceae; Nicotiana.

REFERENCE 1 (bases 1 to 962)
AUTHORS Buell,C.R., Hart,A., Zismann,V., Karamycheva,S.A., Day,B.,
Staskawicz,B., Jin,H. and Baker,B.
TITLE Generation of EST sequences from Nicotiana benthamiana
JOURNAL Unpublished (2003)
COMMENT Other ESTs: EST760389
Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@tigr.org
Clones can be requested from TIGR via potato@tigr.org
Seq primer: ATT TAG GTG ACA CTA TAG.
Location/Qualifiers

FEATURES
source

1..962
/organism="Nicotiana benthamiana"
/mol_type="mRNA"
/db_xref="taxon:4100"
/clone="NBMDS4"
/tissue_type="abiotic and biotic stress-treated leaves,
callus tissue and root tissue"
/lab_host="DH10B-Tona"
/clone_lib="Nicotiana benthamiana mixed tissue cDNA
library, normalized, full-length"
/note="Vector: pCMVSPORT6.1; Site_1: EcoRI; Site_2: NotI;
supplier: RNA was isolated from Nicotiana benthamiana
tissues that include callus, roots from liquid culture
grown plants, heat-stressed leaves (38 C, 3 hr and 6 hr),
cold-stressed leaves (5 C 3 hr, 6hr), and pathogen
challenged leaves (Pseudomonas syringae pv tomato 12 hr;
Xanthomonas campestris pv campestris 12 hr, 18hr;
Pseudomonas syringae pv phaseolicola 18hr, and Xanthomonas
campestris pv vesicatoria 18hr). RNA was isolated from
these tissues and pooled in approximately equal molar
amounts."

ORIGIN

Query Match 23.0%; Score 129.8; DB 14; Length 962;
Best Local Similarity 58.1%; Pred. No. 9e-26;
Matches 270; Conservative 0; Mismatches 187; Indels 8; Gaps 2;

QY 55 CACACCCCTCTGTTTGTAGTGAAGTGAAGCCGAGTGTAAACCCGATTGCCGACTA 114
DB 291 CACACCATTTTCAGTATTTTGTGGGATGAAGTGTATATTAACAAGAAATTGCTGAGTT 350
QY 115 TTTGCCCGCGTGTGTTTAACTTGAAGCTTGGCGGTGGCGGACAGGGGAC 174
DB 291 TTTGCTCGAAGGGGTTAACAATATGATCACTTGTCTGTGGTTGACAA-----GGAT 404
QY 351 TTTGCCGATCAAGTGTGTTAAGTGAAGTGAAGCCGAGTGTAAACCCGATTGCCGACTA 414
DB 175 GTTTCGCCGATCAGCATGTTGTCGGGGGATGAACACATCGAACACTGACCAAG 234
QY 405 AAGCTCTCTTACCATAGTGTCTCTGGAACAGGAAGTCTTGCAACAAGTTGTAGAG 464
DB 415 GAATGG--TGGGGACCCCGGTAAATGTGTAGCAATCTCCAGATTTGGCCAACTTGGC 472
QY 465 CAGCTCAACAAGCTTGGAAGTCTTGAAGTGAAGGATGTGCAAGAGAACCAAGTT 524
DB 295 GAAAGGGAATTGATCTGTGTGAAGTGAAGCCGCAATGCCCTAACCGAGGGAAGTAT 354
QY 525 GAACGAGATTGATGCTCATTAAGCTGAATGACAGTGAAGCAAAAGCTGAATTTATG 584
DB 355 GAGTAGCCCAAGTATTCGGGCGCCGCAATTGATATCTCGAAGACACCGTCCACATC 414
QY 585 TGGTTGTGACATCTTCAGAGGGAATTTGTGATATATCAGACACTTTTGTGACTATC 644
DB 415 GAATGG--TGGGGACCCCGGTAAATGTGTAGCAATCTCCAGATTTGGCCAACTTGGC 472
QY 645 GAGGTGACTGTGACCCGAGGAAATGTGCTGTTTGAAGAAATTTAAGCAAGTTGGG 704
DB 473 ATTAAGAGGTGCTCGAACGGGCAAAATTTGCTTTGGTGGCGGAA 517
QY 705 ATCAAAAGACTAGCCCGAAGCTGAAAGATGCTTTTGAAGAGTGA 749

RESULT 8
 CK243941/c 887 bp mRNA linear EST 12-DEC-2003
 LOCUS
 DEFINITION EST127578 potato callus cDNA library, normalized and full-length
 Solanum tuberosum cDNA clone F0CA635 3' end, mRNA sequence.
 ACCESSION CK243941
 VERSION CK243941.1 GI:39789035
 KEYWORDS
 SOURCE
 ORGANISM
 Solanum tuberosum (potato)
 Solanum tuberosum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 asterids; lamiales; Solanales; Solanaceae; Solanum.
 REFERENCE
 AUTHORS Buell, C.R., Hart, A., Zismann, V., Karanymcheva, S.A. and Baker, B.
 TITLE Generation of ESTs from potato callus tissue
 JOURNAL Unpublished (2003)
 COMMENT Other ESTs: EST127579
 Contact: Robin Buell
 The Institute for Genomic Research
 9712 Medical Center Dr, Rockville, MD 20850, USA
 Email: potato-arr@tigr.org
 Clones can be requested from TIGR via potato@tigr.org
 Seq primer: TTTTCTTTTCTTTTCTTTTCTTTT.
 Location/Qualifiers
 1..887
 /organism="Solanum tuberosum"
 /mol_type="mRNA"
 /culivar="kennebec"
 /db_xref="taxon:4113"
 /clone="F0CA635"
 /tissue_type="callus"
 /lab_host="DH10B-Tona"
 /clone_1ib="potato callus cDNA library, normalized and full-length"
 /note="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI; supplier: RNA was isolated from Solanum tuberosum var. kennebec callus tissue grown on solid media."
 ORIGIN
 Query Match 22.9%; Score 129.4; DB 14; Length 887;
 Best Local Similarity 55.4%; Pred. No. 1.1e-25;
 Matches 271; Conservative 0; Mismatches 216; Indels 2; Gaps 1;
 QY 54 ACAACCTCTCTCTGTTTGAAGTGAAGAAACCGGAGTGTACCCGCAATGCCGACT 113
 DB 809 ACAACATTAAGCATGCTGTGCAACATGCTCTGGGGGCTCTTAATTGATGCTGT 750
 QY 114 ATTGCCCGCGCGTGTATTACATTGAGAGCTTGGCGGTGGCGGAGCAAGCGGGA 173
 DB 749 TATATCTGAGAGGAGTAAATTTTCAGAGCTTGTCTGAGGCTTCAAGACCAAGG 690
 QY 174 CGTTTCCGATCAACATGCTGTGCGGGGAGTGAACACCATGCAACTGACCA 233
 DB 689 ACTTTCAGTATTACCAAGTGTCTCTGGAATGATGAATAGGGAATTGTCA 630
 QY 234 GCAACTCTACAGTGTGTTAACGTAAATTAAAGTACAGACATACCGAATCCCTGT 293
 DB 629 GCAATTTAACAGTTGATAGTGTTCATGAGTTCAAGATTAATCTATTACCAATTTG 570
 QY 294 GAAAGGGAATGATGCTGTGAGAGTGAAGCGCCATGCCCTTACCGAGGAGTGTAT 353
 DB 569 TGACCGGAGCTATGCTCATTAAGTAGCGGCAATGACAGCAAGAGGAGATGTCT 510
 QY 354 TGACTTACCCAGTATTTCCGGGCGGCATTTGAGATCTCCGAGACACCGTCAAT 413
 DB 509 TGAATTTCCAGTATTTTCCGTGTAAACCTGTGATGTGTGATCATACATTAAT 450
 QY 414 CGAATGCTGGGGGACCCGGGTAAATGTAGCAATCTCCAGAT - GTTGGCCAATGG 471
 DB 449 GGAATTAAGTGAAGACTTCAACAAGCTGTGGCATTCACAAGATTAAGAGCTTATG 390
 QY 472 CATTAAGAGGTGCTGAAAGGCAAAATTTGCTTGTGCGGAAATCGGCGTCAATAC 531

DB 389 CATTGTGAGGTGCTGTGACAGGCGGAGTACATTGTGACGAGATCTGTGTGACTC 330
 QY 532 GGAATATCT 540
 DB 329 GACATACCT 321
 RESULT 9
 CA242014
 LOCUS
 DEFINITION SCEPFL3082B07.g Saccharum officinarum FL3 Saccharum officinarum
 cDNA clone SCEPFL3082B07 5', mRNA sequence.
 ACCESSION CA242014
 VERSION CA242014.1 GI:35317200
 KEYWORDS
 SOURCE
 ORGANISM
 Saccharum officinarum
 Saccharum officinarum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; liliopsida; Poales; Poaceae; PACAD
 clade; Panicoideae; Andropogoneae; Saccharum.
 REFERENCE
 AUTHORS Vettore, A.L., da Silva, F.R., Kemper, E.L. and Arruda, P.
 TITLE The libraries that made SUCEST
 JOURNAL Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
 COMMENT Contact: Arruda P
 Centro de Biologia Molecular e Engenharia Genetica
 Universidade Estadual de Campinas
 Caixa Postal 6010, 13083-970, Campinas SP, Brazil
 Tel: 55 19 3788 1137
 Fax: 55 19 3788 1089
 Email: parnuda@unicamp.br
 Clone distribution: clone distribution information can be found
 through the Brazilian Clone Collection Center (BCCC) at
 http://www.bccccenter.fcav.unesp.br
 Plate: 082 row: B column: 07
 Seq primer: 17 Promoter Primer.
 Location/Qualifiers
 1..584
 /organism="Saccharum officinarum"
 /mol_type="mRNA"
 /db_xref="taxon:4547"
 /clone="SCEPFL3082B07"
 /lab_host="DH10B"
 /clone_1ib="Saccharum officinarum FL3"
 /note="Organ: Base of developing inflorescence (5cm-long); Vector: pSPORT1; Site 1: SalI; Site 2: NotI; An unidirectional cDNA library generated from (base of developing inflorescence (5cm-long)). cDNA was prepared from polyA+ mRNA using Superscript Plasmid System Kit (Invitrogen). The double-strand cDNAs were fractionated in a sepharose CL-2B 40cm-columns and fragments sizing between 0.8 and 1.5 kb were directionally cloned into the vector. Details of each source of RNA and library construction can be obtained at
 http://succest.iad.ic.unicamp.br/public"
 ORIGIN
 Query Match 22.8%; Score 129; DB 14; Length 584;
 Best Local Similarity 58.0%; Pred. No. 1.2e-25;
 Matches 265; Conservative 0; Mismatches 190; Indels 2; Gaps 2;
 QY 86 CCGAGTCTTAACCCGATTCGCCGACATTTGGCCCGCGTGTTAACATTGAGACT 145
 DB 9 CTGGTCTTCAACATTGTAAACAGAGCTTTGCTGCGAGGCTACAAATACAGAGCC 68
 QY 146 TGGCGTGGGGTGGCGGAGACAGGGGACGTTTCCCGCATCACCATGATGATGCGGGG 205
 DB 69 TTGCTGTGGCCAGCTGAGAGAGGACATTTGCGGTATTAACAAGTGTCTCTGTA 128
 QY 206 ATGAGAACACATGAGACATGACCAAGCAACTCTTCAAGTTGTTTAACTAATTAAG 265
 DB 129 CTGTTGATTCATTGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 188

OY 266 TACAGACATCACCGAACTCCCTGTGTGAAAAGGAATTGATGCTGTGAGTGACG 325
DB 189 TTGATGACATTAACCCCTCACTTTTGTGTGAAAAGGAATGATGCTTAATTAAGTTTCTG 248
OY 326 CCAATGCCCTTAACCGAGCGGAAGTATGATGATGAGTCCAGGATTTCCGGCCCGCATTTG 385
DB 249 TAAACACTGCTGTCTGGAGGGAATCTAGATATTTGCCGAATCTCCGACAAACCTG 308
OY 386 TGGATATCTCCGAGAGACACCGTAC-CATCGAATGTGGGGGACCCGGGTAAATGTAG 444
DB 309 TTGATGTTTCTGACATTAAGTAAAGCTTCACTTACGAGATTTTGACAGATGTTG 368
OY 445 CAATCTCCAGATGTGG-CCAAGTTGGCATTTAAAGAGTGGCTCGAAGCGGCAAAATTG 503
DB 369 CACTACAAAGGTATTTGAGGACATATGAGCATCTCGAGGTGCGCAGAACTGACGAGTGG 428
OY 504 CTTTGTGGGGGAATCCGGCGTCAATAGGAATATCT 540
DB 429 CACTGTCTGTGAATCGAAGGTGCACTCCAGTACT 465

RESULT 10
CB672821/c 626 bp mRNA linear EST 09-APR-2003

LOCUS CB672821
DEFINITION Oryza sativa (japonica cultivar-group) cDNA
clone OSUNBE07B04 3', mRNA sequence.

ACCESSION CB672821
VERSION CB672821
KEYWORDS EST.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Erihartoideae; Oryzaceae; Oryza.

REFERENCE 1 (bases 1 to 626)
Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,D., Zhou,B., Mazur,E.,
Kudrna,D., Dean,R., Soderlund,C., Wang,R. and Wang,G.
Large-scale identification of ESTs involved in the interaction
between rice and Magnaporthe grisea
Unpublished (2003)
JOURNAL Contact: Rod Wing
COMMENT Arizona Genomics Institute
University of Arizona
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
85721-0088, USA
Tel: 520 626 3967
Fax: 520 621 9288
Email: http://genome.arizona.edu
PCR Primers
FORWARD: gta aac cga cgg cca gtc
BACKWARD: gga aac agc tat gac cat g
Plate: 07
Row: B
Column: 04
Seq primer: gga aac agc tat gac cat g.

FEATURES
source
1..626
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="OSUNBE07B04"
/tissue_type="Leaf"
/dev_stage="3 week"
/lab_host="DH10B"
/clone_1lb="OSUNBE"
/note="Vector: pBluescript II KS +; site 1: EcoRI; site 2:
XhoI; 24 hrs after inoculation with Rice Blast (70-15)F"

ORIGIN

Query Match 22.8%; Score 129; DB 14; Length 626;
Best Local Similarity 56.6%; Pred. No. 1.2e-25;
Matches 277; Conservative 0; Mismatches 210; Indels 2; Gaps 2;

OY 54 ACACACCTCTCTGTTTATGTTAAAGTAGAGCCGAGTGTCTAACCCGATTGCCGACT 113
DB 623 ACATACCTATTCATCCCTGTGCAATGATTTGCCCTGTGTTTCAACATTGTTACAGGGAT 564
OY 114 ATTTGCCGCGGAGTTTAAACATTGAGAGCTTGGCGGAGGAGTGGCGGAGAGAGAGG 173
DB 563 CTTTGTGCGAAGGCTACATATACAGAGTTTGTGTAGGCCACAGCTGAAATGTCAGG 504
OY 174 CGTTTCCGCAATCACCATGTTGTGTCGGGGAGTAGAAACACATCGAACAATGACCA 233
DB 503 CTTTGGCTATTAACAACAGTTGCTCTGGAACAAGATGATCATTTGAAATTTAGTTCA 444
OY 234 GCAACTCTCAAGTTGTTAAGTAAATTAAGTACAGGACATCACGAAATCCCTGTGT 293
DB 443 GCACTTAACAACCTTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 384
OY 294 GGAAGGAAATGATGCTGTGTAAGTGAAGTGAAGCCCAATGCCCTTAACCGAGGAAGTAT 353
DB 383 TGAAGAGAACTATGCTTATCAAGTTTCTGTGAACAATGCTGCTCGAGAGACATACT 324
OY 354 TGAAGTACCGCAGATTTTCCGGCCCGCATTTGTGATATCTCCGAAGACACCGTACCAT 413
DB 323 AGATATTGCTGAATCTTCCGGGCAAAATCTGTGATGTTGTGATGATGATGATGATGAT 264
OY 414 -CGAATGTGGGGGAGCCCGGATAAATGTAGCAATCTCCAGATGTTGG-CCAAATTGG 471
DB 263 ACAGCTTACTGGGAGATCTGACAAAGATGTTGATTAACAAGCTGTTGAGACCTTAATG 204
OY 472 CATTAAGAGGTGGCTCGAAGCGGCAAAATGCTTTGTGCGGAAATCCGGCTCAATTC 531
DB 203 CACTGTGAGGTGCGCAGAAACAGGGCGAGTGGCGCTGTCCGAGATCGGTGTGATTC 144
OY 532 GGAATATCT 540
DB 143 CAAGTACT 135

RESULT 11
BP184863
LOCUS BP184863
DEFINITION BP184863 pms rice callus cDNA Oryza sativa (japonica
cultivar-group) cDNA, mRNA sequence.
ACCESSION BP184863
VERSION BP184863
KEYWORDS EST.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Erihartoideae; Oryzaceae; Oryza.
REFERENCE 1 (bases 1 to 695)
Moriguchi,K., Ito,Y., Yamazaki,Y. and Kurata,N.
Finding of various plant nuclear proteins using yeast nuclear
transportation trap system - a proteomal approach
Unpublished (2003)
JOURNAL Contact: Kazuki Moriguchi
COMMENT Plant Genetics
National Institute of Genetics
Yata 1111, Mishima, Shizuoka 411-8540, Japan
Tel: 81-55-981-6872
Fax: 81-55-981-6879
Email: kmoriguc@lab.nig.ac.jp
cDNA clone obtained from nuclear transportation trap system
encoding a protein similar to Arabidopsis thaliana acetolactate
synthase-like protein.
location/Qualifiers
1..695
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/db_xref="taxon:39947"
/tissue_type="callus"
/dev_stage="4 days after regeneration treatment"
/clone_1lb="pMS rice callus cDNA"

FEATURES
source

Query Match	22.8%;	Score 129;	DB 12;	Length 695;
Best Local Similarity	56.6%;	Pred. No. 1.3e-25;		
Matches 277; Conservative	0;	Mismatches 210;	Indels 2;	Gaps 2

QY	54	ACACACCCCTCTGTTTATGTTGAAGAGGAAGCCGAGTGTAAACCCGATTCGCGAAT	113
Db	62	ACAATACTATTCATCCTCTTGTCATAGATTCGCCCTGGTCTTCAACATTTGTTACAGGGGT	121
QY	114	ATTTCGCCGCCCTGCTTTTAAACATTGAGAGCTTGGCGGTGGGGTCCGGCAACAGGGGGA	173
Db	122	CTTTGTCTCGAAGGCTACAAATATACAGAGTCTTGTGTAGGCCACGCTGAAAGTCAAG	181
QY	174	CGTTTCCCGCATCATCCATGTGTGTGTCGCCGGGAGATGAGAAACCATGTGAAACAATGACCA	233
Db	182	CCTTTGCGGTATTAACAACAGTTTGCCTCCGGAAACAGATGAAATTCATTGAGAGTTGATTCA	241
QY	234	GCAACTCTTACAAAGTTGGTTTAAAGTAAATTAAGTACAGAGATCACCGAAACTCCCTGTGT	293
Db	242	GCAAGCTTAACAAACTGTTGATGTGATGTCAGAGTTCAAGATATATACTCATTCGCTTTTGC	301
QY	294	GGAAAGGGGAATTGATGCTGTGTGAAAGGTGAGAGGCGCAATGCCCTTAACCGAGAGAACTAT	353
Db	302	TGAAAGAGAACTTAATGCTTATCAAGGTTTCTGTGACACTGCTGCTCGGAGAGACATACT	361
QY	354	TGAGCTAAGCCACGAGTATTCGCGAGCCGCAATTTGTGATATCTCCGAAGACACGTCACAT	413
Db	362	AGATATTTGCTGAAATCTTCCGGGGCAAAATCTGTGATGTCTTGATCACACTGTTAAGTT	421
QY	414	-GGAAATGATGGGGGACCCCGGGTAAATAATGTGAGCAATCTCTCAATGTTGG-CCAAAGTTGG	471
Db	422	AACAGCTTACTGGGGGATCTCGACAAAGATGTGTGACTTACAAAGGCTTTTGAACCTTAATGG	481
QY	472	CATTAAAGAGTGTGCTCGAACGCGCAAAATTTGTTTGGTCCGGGAATCCGGGCTCAATAC	531
Db	482	CATCTGTGAGGTGCGCAGAAACAGGGCGAGTGGGCGCTGTGTCCGGAAATCCGGTGTGATTC	541
QY	532	GGAAATATCT	540
Db	542	CAAGTAACT	550

FEATURES	source
RESULT 12	
CF210202	
LOCUS	697 bp mRNA linear EST 01-AUG-2003
DEFINITION	CAB20006_11ta_Fa_C06 Cabernet Sauvignon Flower bloom - CAB2 Vitis
ACCESSION	Vinflera CDNA clone CAB20006_11ta_Fa_C06 5', mRNA sequence.
VERSION	CF210202
KEYWORDS	CF210202.1 GI:33404575
SOURCE	EST
ORGANISM	Vitis vinifera
REFERENCE	Vitis vinifera
AUTHORS	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; Vitaceae; Vitis.
JOURNAL	1 (bases 1 to 697)
COMMENT	Goes da Silva, F., Iandolino, A., Lim, H., Baek, J., Jones, K. and Cook, D.
	Expressed sequence tags from Vitis vinifera 'Cabernet Sauvignon' berries at various developmental stages
	Unpublished (2003)
	Contact: Douglas Cook, PhD
	CABs Genome Facility
	UC Davis, Plant Pathology
	One Shields Ave., Davis, CA 95616, USA
	Tel: 530 754 6561
	Fax: 530 754 6617
	Email: drcocok@ucdavis.edu
	Seq primer: ACGGTACCGACATATGCC.
	Location/Qualifiers
	1..697
	/organism="Vitis vinifera"

```

/mol_type="cRNA"
/cultivar="Cabernet Sauvignon"
/db_xref="taxon:29760"
/clone="CAB20006_title_Fa_C06"
/sex="Hermaphrodite"
/dev_stage="Bloom"
/clone_1ib="Cabernet Sauvignon Flower bloom - CAB2"
/notes"Organ: Flower - Bloom; Vector: pDNR; Site 1: 5fl;
Site 2: 5fl; CAB2 is a cDNA library of Vitis vinifera cv
'cabernet sauvignon' clone 8 berries. Samples were
collected at full bloom (80 to 100% flowers showing
deniscence of calyptas or caps and anthers fully
extended). Sampled vines were located at the University of
California, Davis, Experimental Vineyard. cDNAs were made
by oligo-dt priming and directionally cloned. 5' and 3'
adaptors were used in cloning as follows:
5'-AAGCAGGTGATTCACGACGAGGTGCATCATGACGCCGG-3' and
5'-ATTTCAGAGCGCAGCGCGCGCGACATG-(30)NN-3'. Library was
constructed using the Clontech Creator SMART Kit and
size-selected to contain the 0.5-3 kb size fraction."

```

Query Match	22.8%	Score 128.8	DB 14	Length 697
Best Local Similarity	55.3%	Pred. No. 1.5e-25		
Matches 270	Conservative 0	Mismatches 217	Indels 1	Gaps 1
QY	54	ACACACCCCTCTCTGTTTAACTTGAAGTGAAGCCGAGAGTGTCAACCCGATTTGCCGACT	113	
Db	45	ACACACTATATCATGTTGTAATAATATATGCTCCGAGATTTCAATCTTGTACTAGGGT	104	
QY	114	ATTTCGCCGCCGTGTTTAACTTTGACAGCTTGCCTGCGGTCCGCCGAACGAGGGA	173	
Db	105	TTTTCGAAGAGGGGATTAACATCCAGAGCTGCGAGTTGTCATGCAGAAATTGAGGG	164	
QY	174	CGTTTCCGCACTACCATGTGTGTCGCGGGGATGAGAACCATGTGACAACTGACCA	233	
Db	165	TCTCTCTGCATTACACTGTGTTCTGTGTAAGATTAATCATTAACAGATTGATGCA	224	
QY	234	GCAACTCTACAAGTGTGTTAACGTAAATTAAGTAGACGACATCACCGAACTCCCTGT	293	
Db	225	ACAACCTTAAGAAAGCTAATAGATCTTCATGATGTTCAGAAATTAACCATTAACCTTTC	284	
QY	294	GGAAAGGGAATTGATGCTGTGGAAGGTGACGCCAATGCCCTAACCGACCGAATGAT	353	
Db	285	TGAGCGAGAGTTGATGTTGAATAAGATGCTGTGAATGCTGTGCTGCCGCTGAGTGCT	344	
QY	354	TGAGCTAGCCGAGGATTCGCGGCCCGCATGTGTGATCTCCGAAACCGTCATCAT	413	
Db	345	TGATTTTCCGACATTTTTCAGGGCCAAAGCCATTGACGTGTGATCACACATTAACCT	404	
QY	414	CGA-ATGTGTGGGGGACCCGGGTAAATGTGTAGCAATCTCCAGATGTTGGCCCAATTGGC	472	
Db	405	TGAGCTTACAGAGCTTTTACGAAAAAAGTTGTCTGCAAAAGTTGTATGAGCCCTATGG	464	
QY	473	ATTAAAGAGGTGCTCGAACGGGCAAAATTTGCTTTGTGTGGGAAATCCGGCGCTCAATCG	532	
Db	465	CATTGTGAGGTGACGCAACTGGCGGGTGGCTGTGTGCTGTGAGTCGGGTGTGATCTCC	524	
QY	533	GAAATCT 540		
Db	525	AATTAATCT 532		

RESULT 13	CD429148	548 bp	mRNA	linear	EST 03-JUN-2003
LOCUS	CD429148				
DEFINITION	CD429148				
	ETH1_2.A06.g1.A002 Ethylene-treated seedlings Sorghum bicolor CDNA				
	clone ETH1_2.A06.A002 5', mRNA sequence.				
ACCESSION	CD429148				
VERSION	CD429148				
KEYWORDS	CD429148.1 GI:31344791				
SOURCE	EST				
	Sorghum bicolor (Sorghum)				
ORGANISM	Sorghum bicolor				

REFERENCE
AUTHORS
Kuek, P., Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD
clade; Panicoidae; Andropogoneae; Sorghum.
1 (bases 1 to 548)
Klein, R.R., Liang, C., Sun, F., Sullivan, R., Shah, M., Salzman, R.,
Chen, J., Gonzalez, W., Lane, S., Miller, W., Nanda, P.,
Olaseinde, O., Eastman, A. and Pratt, L.H.
An EST database from Sorghum: 1-aminocyclopropane-1-carboxylic acid
(ACC)-treated seedlings
Unpublished (2003)

JOURNAL
COMMENT
Contact: Cordomier-Pratt MM
Laboratory for Genomics and Bioinformatics
The University of Georgia, Department of Plant Biology
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 583 0210
Email: mmpratt@uga.edu

FEATURES
source
1. 548
/organism="Sorghum bicolor"
/mol_type="mRNA"
/cultivar="BTX623"
/db_xref="taxon:458"
/clone="ETH1_2_A06_A002"
/lab_host="DH10B-T1 phage-resistant E. coli"
/note="Vector: pME18S-Fl3; Site 1: XhoI; Site 2: XhoI; The
library was prepared from polyA+ RNA from seedlings grown
in hydroponic culture. At 8 days of age, medium was
supplemented with 0.5 mM 1-aminocyclopropane-1-carboxylic
acid (ACC) to induce endogenous ethylene (ETH) production.
Roots and shoots were harvested after 27 and 72 hr and
material from both time points was combined prior to RNA
isolation. Double-stranded cDNA was cloned
unidirectionally into different DraIII sites of the
pME18S-Fl3 vector (5-prime DraIII site is CAGCTGTCG,
3-prime DraIII site is CACCATGTG). XhoI excises the cDNA
insert."

ORIGIN

Query Match 22.8%; Score 128.6; DB 14; Length 548;
Best Local Similarity 57.7%; Pred. No. 1.5e-25;
Matches 267; Conservative 0; Mismatches 194; Indels 2; Gaps 2;
55 CACACCTCTCTGTTTAACTGAAGTGAAGCCGAGTCTAACCCGATTGCCGAGCTA 114
86 CATACCTCTCTCAATCTTGTAATGACTGCTGCTGCTCAACATTTGTAACAGAGATC 145
115 TTGGCCCGCGGTTTAACTGAAGTGAAGCCGAGTCTTGCGGAGTGAAGAGGAG 174
146 TTGGCTCGAGGGGCTACATTAACAGACCTTCTGTTGGCCGAGTGAAGAGAGC 205
175 GTTTCGCCGATCAACAGTGTGTCGGGAGTGAAGAACCCATGCAACTGCAAG 234
206 ATTGGCGATTACCAAGTGTCTCTGTAAGTGAAGTGAAGTGAAGTGAAGTGAAG 265
235 CACCTCTCAAGTGTGTTAACTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 294
266 CAGCTTAAAGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 325
295 GAAAGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 354
326 GAAAGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 385

Query 355 GAGCTAGCCAGGATATTCGGGGCCGAGTGGATATTCGGAAGACCCGTCAC-CAT 413
386 GATATTCGGAATCTTCGAGCAAAACCTGTTGATGATGATGATGATGATGATGAT 445
414 CGAATGTGGGGGACCCGGTAAATGATGATGATGATGATGATGATGATGATGATG 472
446 CAGCTTACGAGATTTTGAAGATGATGATGATGATGATGATGATGATGATGATG 505
473 ATTAAGAGTGGCTGGAACCGGCAAAATTTGCTTTGTCGCG 515
506 ATCTGGAGGTGCGCAGAACTGAGACAGTGGACCTGTCGCG 548

RESULT 14
BG581436
LOCUS
DEFINITION
EST483170 GVN Medicago truncatula cDNA clone gvn-64H22 5' end,
mRNA sequence.
ACCESSION
BG581436
VERSION
BG581436.1 GI:13596500
KEYWORDS
EST.
SOURCE
Medicago truncatula (barrel medic)
ORGANISM
Medicago truncatula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifoliaceae;
Medicago.

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
1 (bases 1 to 782)
Fedorova, M., Peterson, B.L., Samac, D.A., Vance, C.P., Gant, G.S.,
Town, C.D., Van Aken, S., Utecher, T., Cho, J., and Fraser, C.M.
ESTs from one month old nitrogen-fixing root nodules of Medicago
truncatula, 2001
Unpublished (2001)
Contact: Carroll P. Vance
Department of Agronomy and Plant Genetics
University of Minnesota
411 Borlaug Hall, 1991 Upper Buford Circle, St. Paul, MN 55108 USA
Tel: 612 625 5715
Fax: 651-649-5058
Email: vance004@maroon.tc.umn.edu
University of Minnesota name: M382496e TIGR sequence name:
M382496e More information is available at: <http://www.medicago.org>
Seq primer: SKmod (CTA GAA CTA gta gat cc).
Location/Qualifiers
1. 782
/organism="Medicago truncatula"
/mol_type="mRNA"
/cultivar="genotype A17"
/db_xref="taxon:3880"
/clone="pGVN-64H22"
/tissue type="N2-fixing root nodules"
/dev stage="effective root nodules harvested one month
post inoculation with Sinorhizobium meliloti"
/lab host="E. coli strain X10LR"
/clone lib="GVN"
/note="Vector: pBluescript SK-; Site 1: EcoRI; Site 2:
XhoI; cDNA was prepared from polyA+ enriched RNA from
effective root nodules harvested one month post
inoculation with Sinorhizobium meliloti. The cDNA was
directionally ligated into the uni-ZAP XR vector from
Stratagene and packaged using GigaPack III Gold packaging
extracts. Plasmids containing cDNA inserts were excised
from the recombinant lambda-ZAP phage using Ex-Assist
helper phage and propagated in X10LR cells."

FEATURES

source
1. 782
/organism="Medicago truncatula"
/mol_type="mRNA"
/cultivar="genotype A17"
/db_xref="taxon:3880"
/clone="pGVN-64H22"
/tissue type="N2-fixing root nodules"
/dev stage="effective root nodules harvested one month
post inoculation with Sinorhizobium meliloti"
/lab host="E. coli strain X10LR"
/clone lib="GVN"
/note="Vector: pBluescript SK-; Site 1: EcoRI; Site 2:
XhoI; cDNA was prepared from polyA+ enriched RNA from
effective root nodules harvested one month post
inoculation with Sinorhizobium meliloti. The cDNA was
directionally ligated into the uni-ZAP XR vector from
Stratagene and packaged using GigaPack III Gold packaging
extracts. Plasmids containing cDNA inserts were excised
from the recombinant lambda-ZAP phage using Ex-Assist
helper phage and propagated in X10LR cells."

Query Match 22.8%; Score 128.6; DB 12; Length 782;
Best Local Similarity 59.1%; Pred. No. 1.5e-25;
Matches 277; Conservative 0; Mismatches 184; Indels 8; Gaps 3;
51 GAAACACACCTCTCTGTTTAACTGAAGTGAAGCCGAGTGAAGTGAAGTGAAGTGAAG 110

Db 228 GAGACACAGATTGGTGGTTCGTTGGTGGATGAGATGAGATTAACCGAATAGCTGG 287
 QY 111 ACTATTTGCCCCCGCTGGTTTAACTATGAGAGCTGGCGGTGGGGTGGCGAAGAGG 170
 Db 288 AGTGTTCGGAAGAAAGAGATACACATTGAGTCATGAGCTGTGGTTGAATCAAGTAG 347
 QY 171 GAGACCTTTCCCGCATCACCATTGTTGGTCCCGGGGAGTGAACAACCATGCAACTGAC 230
 Db 348 GGCCTCTT-----TCACCATGTGTTTCTGTTAGTACGATAGGGTTTACGTCAAGTTGT 401
 QY 231 CAAGCACTCTCAACAGTTGGTTAAGTAAATTAAGTACAGGACATCAACCGAAATCCCTG 290
 Db 402 TGAACAGCTTCAGAAAGCTGTGCAATGTTTGAAGTGAAGATCTTTCAGCTGACCTCA 461
 QY 291 TGTGAAAAGGAATTGATGCTGTGTAAGGTGAGCGCCCAATGCCCTTAACGAGCGAAGT 350
 Db 462 AGTAAACGTGAATGATGCTGTGAAAGTTCATGCAAGATCCCAAAACGCTGTGAGAT 521
 QY 351 GATTAGCTAGCCCAAGTATCCCGGGCCCGCATTTGGTGAATATCTCCCAAGACACCCGAC 410
 Db 522 TAAAGTGTGGTGCACACCTTCAGAGCCCAAGATGCTGGAACATCTCGGAACATTCGTCGAC 581
 QY 411 CATCGA-ATGGTGGGGGACCCCGGTAAATGATGCAATCC-TCAGATGTTGGCCAGT 468
 Db 582 GATTGAGGTACTGGAGAGCCAGGGAAGATGTTGCACTCCAGAGAAATTTCCGCAAGTT 641
 QY 469 TGGCATTAAGAGGTGGCTCGAACGGGCAAAATTCCTTGGTGGCGGAA 517
 Db 642 TGGAAATTAAGAAATAGCTAGAACCCGGAAGATTGCAATTAAGAAAGGAA 690

RESULT 15
 CK243942 891 bp mRNA linear EST 12-DEC-2003
 DEFINITION EST1727579 potato callus cDNA library, normalized and full-length
 Solanum tuberosum cDNA clone POCa635 5' end, mRNA sequence.
 ACCESSION CK243942
 VERSION CK243942.1 GI:39789037
 KEYWORDS EST.
 SOURCE Solanum tuberosum (potato)
 ORGANISM Solanum tuberosum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 asterids; lamids; Solanales; Solanaceae; Solanum.
 REFERENCE 1 (bases 1 to 891)
 AUTHORS Buell, C.R., Hart, A., Zismann, V., Karamycheva, S.A. and Baker, B.
 TITLE Generation of ESTs from potato callus tissue
 JOURNAL Unpublished (2003)
 COMMENT Other ESTs: EST1727578
 Contact: Robin Buell
 The Institute for Genomic Research
 9712 Medical Center Dr, Rockville, MD 20850, USA
 Email: potato-array@igr.org
 Clones can be requested from TIGR via potato@igr.org
 Seq primer: ATT TAG GTG ACA CTA TAG.

FEATURES
 source
 1..891
 Location/Qualifiers
 /organism="Solanum tuberosum"
 /mol_type="mRNA"
 /cultiyar="Kennebec"
 /db_xref="taxon:4113"
 /clone="POCa635"
 /tissue_type="callus"
 /lab_host="DH10B-Tona"
 /clone_id="potato callus cDNA library, normalized and full-length"
 /note="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI; supplier: RNA was isolated from Solanum tuberosum var. Kennebec callus tissue grown on solid media."

ORIGIN

Query Match 22.7%; Score 128.2; DB 14; Length 891;
 Best Local Similarity 57.8%; Pred. No. 2.5e-25;

Matches 269; Conservative 0; Mismatches 188; Indels 8; Gaps 2;
 QY 55 CACACCTCTGCTGTTTAACTATGAGATGAGCCGAGAGCTAACCCGATTCGCGACTA 114
 Db 310 CACACATTTCAATATTTGTTGGGATGAAGATGCAATATATATAGATATGCTGAGTT 369
 QY 115 TTTGCCCCGCTGGTTTAACTGAGAGCTTGGCGGTGGGCTGGCGGAAAGGGGAGC 174
 Db 370 TTTGCTCGAAGGGGTTTACAAATTTAGTACACTTCTGTTGGTTGAACAGGATTAAGCT 429
 QY 175 GTTTCGCCGATCAGCATGCTGCTGCGGGGAGTGAACACCATGGAACAACGCAAG 234
 Db 430 CTTT-----TTACATAGTGTCTCTGGAAGTGAAGGTTTCAACCAAGTTGAGG 483
 QY 235 CAATCTACAGTTGTTTAACTGTAATTAAGTACAGACATCAACGAAATCCTGTGTG 294
 Db 484 CAGCTCAACAGCTTGTGAATGTCGGAAGGTGAGGATCTGTCAAGAGAACACACAGTG 543
 QY 295 GAAAGGATTTATGCTGTGAGTGAAGTGAAGCCCAATGCCCTTAACCGAGCGGAAGTAT 354
 Db 544 GAACGAGAAATGATGCTCATAAACTGAATGCAATGCAACAAACAAAGCTGAATCATG 603
 QY 355 GAGTAGCCCAAGTATTCGCGGCCCGCATTTGTGATATCTCCAGAGACACCGTCACATC 414
 Db 604 TGGTTGTGGAATCTTCAGAGAGAAATTTGTGATATATAGAACACTTTCTGACTATT 663
 QY 415 GA-ATGGTGGGGGACCCGGGTAAATGATGCAATCTCCAGATGTTGGCCAAAGTTGCC 472
 Db 664 GAGGTAACTGTGATCCAGGGGAAATGCTGCTGTTCTGAGAAATTTTAAGCAAGTTGGG 723
 QY 473 ATTAAGAGGTGGCTCGAACGGGCAAAATTCCTTGGTGGCGGAA 517
 Db 724 ATTAGCGAACTAGCCCGAACTGGGAAAGATTGCTTTGAGACAGAA 768

Search completed: July 25, 2004, 07:28:13
 Job time : 1616.95 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 25, 2004, 04:04:07 ; Search time 2268.95 Seconds

(without alignments)
10792.997 Million cell updates/sec

Title: US-09-893-033-17

Sequence: 1 GTGGAATTCACCAATG.....CCCTGATCCCAAGTTTtag 565

Scoring table:

IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 3470272 segs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seg length: 0
Maximum DB seg length: 2000000000Post-Processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_hg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_hum:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pac:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_hg_hum:*
31: em_hg_inv:*
32: em_hg_other:*
33: em_hg_mus:*
34: em_hg_pln:*
35: em_hg_rnd:*
36: em_hg_tam:*
37: em_hg_vrt:*
38: em_sy:*
39: em_hgo_hum:*
40: em_hgo_mus:*
41: em_hgo_other:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	543	96.1 105795	1 SYCCPNC	D64001 Synechocyst
2	272.2	48.2 342850	1 AP003597	AP003597 Nostoc sp
3	258.8	45.8 300450	1 AP005371	AP005371 Thermosyn
4	235.6	41.7 293350	1 AP006572	AP006572 Gloeobact
5	211	37.3 349652	1 BX569690	BX569690 Synechoco
6	207.8	36.8 7550	8 D63676	D63676 Cyanidium c
7	195	34.5 8200	8 D63675	D63675 Cyanidiesch
8	195	34.5 149987	8 AB002583	AB002583 Cyanidios
9	179.6	31.8 191028	8 PP038804	PP038804 Porphyra pu
10	177	31.3 6982	8 AF233069	AF233069 Galdieria
11	176.4	31.2 346683	1 BX572098	BX572098 Prochloro
12	173.2	30.7 164921	8 AF022186	AF022186 Cyanidium
13	161.8	28.6 302399	1 AE017164	AE017164 Prochloro
14	159.6	28.2 121524	8 AF041468	AF041468 Guillardia
15	152.4	27.0 559	1 AP175526	AP175526 Streptomy
16	146.2	25.9 348986	1 BX572093	BX572093 Prochloro
17	141.8	25.1 4210	1 MSGILVB	L49392 Mycobacteri
18	139.4	24.7 308050	1 SC0939124	AL939124 Streptomy
19	138.4	24.5 15939	1 AE007128	AE007128 Mycobacte
20	138.4	24.5 24613	1 AY116644	AY116644 Streptomy
21	138.4	24.5 318050	1 BX248344	BX248344 Mycobacte
22	138.4	24.5 348676	15 BX842581	BX842581 Mycobacte
23	137.2	24.3 11759	1 AE011091	AE011091 Methanosa
24	133.8	23.7 11932	1 AE002508	AE002508 Neisseria
25	133.6	23.7 326301	1 NMA622491	AL162757 Neisseria
26	133.8	23.7 349980	6 AX044033	AX044033 Sequence
27	133	23.5 300272	1 AE017213	AE017213 Geobacter
28	131.4	23.3 347137	1 BX640448	BX640448 Bordetell
29	131.4	23.3 349028	1 BX640433	BX640433 Bordetell
30	129.8	23.0 349305	1 BX640433	BX640433 Bordetell
31	129.6	22.9 281450	1 AP005032	AP005032 Streptomy
32	129	22.8 1883	8 AK058418	AK058418 Oryza sat
33	128.4	22.7 11318	1 AE013291	AE013291 Methanosa
34	126.6	22.4 10818	1 AE015669	AE015669 Shewanell
35	126.4	22.4 570	1 AY142799	AY142799 Helicobact
36	122.4	21.7 300774	1 AE016912	AE016912 Chromobac
37	121.4	21.5 10625	1 AE001995	AE001995 Deinococc
38	120	21.2 13104	1 AE000985	AE000985 Archaeogyl
39	118.4	21.0 200050	1 AL646068	AL646068 Ralstonia
40	117.4	20.8 110000	2 LMFCHR34_10	Continuation (11 o
41	117.2	20.7 44882	1 LMFCHR37	299263 Mycobacteri
42	117.2	20.7 348950	1 MLERPRTV7	AL583923 Mycobacte
43	115	20.4 303050	1 BX321860	BX321860 Nitrobac
44	114	20.2 13257	1 AE014648	AE014648 Bifidobac
45	114	20.2 349980	6 AX492782	AX492782 Sequence

ALIGNMENTS

RESULT 1
SYCCPNC/LOCUS 105795 bp DNA linear BCT 04-JUL-2001
DEFINITION Synechocystis sp. PCC 6803 DNA, complete genome, section:20/27,
2539000-2644794
ACCESSION D64001 AB001338 BA000022
VERSION D64001.1 GI:1001102
KEYWORDS
SOURCE
ORGANISM Synechocystis sp. PCC 6803
Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
REFERENCE
1 Kaneko, T., Tanaka, A., Sato, S., Kotani, H., Sazuka, T., Miyajima, N.,
Sugiura, M. and Tabata, S.
Sequence analysis of the genome of the unicellular cyanobacterium

QY	420	GTGGGGGACCCGGGCTAAATGGTGAACATCCTCAGATGTGGCCAGTT-GGCATTAA	478		
Db	36491	GTGGGGGACCCGGGCTAAATGGTGAACATCCTCAGATGTGGCCAGTTGGCATTAA	36433		
QY	479	GAGTGGCTTCGAACGGGCAAAATGCTTTGGTGGGAAATCCGGGCTCAATAACGAATAT	538		
Db	36431	GAGTGGCTTCGAACGGGCAAAATGCTTTGGTGGGAAATCCGGGCTCAATACGAATAT	36372		
QY	539	CTGAATCCCTGGATCCAGTTTGTG	565		
Db	36371	CTGAATCCCTGGATCCAGTTTGTG	36345		
RESULT 2	AP003597	342850 bp	DNA	linear	BCT 28-NOV-2001
LOCUS	AP003597				
DEFINITION	Nostoc sp. PCC 7120 DNA, complete genome, section 17/19.				
ACCESSION	AP003597	BA000019			
VERSION	AP003597.1	GI:17133729			
KEYWORDS					
SOURCE ORGANISM	Nostoc sp. PCC 7120				
REFERENCE	Nostoc sp. PCC 7120				
AUTHORS	Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.				
TITLE	1. Kaneo, T., Nakamura, Y., Wolk, C. P., Kuritz, T., Sasamoto, S., Matanabe, A., Iriuguchi, M., Ishikawa, A., Kawashima, K., Kimura, T., Kishida, Y., Kohara, M., Matsumoto, M., Matsuno, A., Muraki, A., Nakazaki, N., Shimo, S., Sugimoto, M., Takazawa, M., Yamada, M., Yasuda, M. and Tabata, S.				
JOURNAL	Complete genomic sequence of the filamentous nitrogen-fixing cyanobacterium Anabaena sp. strain PCC 7120				
MEDLINE	DNA Res. 8 (5), 205-213 (2001)				
REFERENCE	21595285				
AUTHORS	2 (bases 1 to 342850)				
TITLE	Kaneo, T.				
JOURNAL	Direct Submission				
REFERENCE	Submitted (02-MAY-2001) Takakazu Kaneo, Kazusa DNA Research Institute, The First Laboratory for Plant Gene Research, Yana 1532-3, Kiseitazu, Chiba 292-0812, Japan				
AUTHORS	(E-mail: kaneo@kazusa.or.jp, URL: http://www.kazusa.or.jp/cyanobase/, Tel: 81-438-52-3935 (ex.238), Fax: 81-438-52-3934)				
FEATURES	Location/Qualifiers				
source	1. 342850				
gene	/organism="Nostoc sp. PCC 7120"				
CDS	/mol_type="genomic DNA"				
gene	/db_xref="taxon:103690"				
CDS	/note="synonym: Anabaena sp. PCC7120"				
gene	/complement(174..1478)				
CDS	/gene="a114593"				
gene	/complement(174..1478)				
CDS	/gene="a114593"				
gene	/note="ORF ID: a114593				
CDS	/hypothetical protein"				
gene	/codon_start=1				
CDS	/transl_table=1				
gene	/protein_id="BAB76292.1"				
CDS	/db_xref="GI:17133730"				
gene	/translation="MDKSTLCVDELSPNIVNVKVLNLDVYAPGOMNIVGDMTIR				
CDS	ATGTGRTAVITOKIRDRATSLYNDPQGGQSAVKYQITDKADITMATLALANKGK				
gene	IGELPLSLNITKADVTSIDLSKIALEIVFCKLNGIPQNPQEPFANSLANNQDA				
CDS	SLIRWALVCTDGLPLDGPFLSKIQGIIISGASNSITQNPVFLVNSLPQSNPTDK				
gene	LEFINGFAVAGMMNLTAKGIRTPQSSISHTGNFIQTADNIDFVAFLDQTNF				
CDS	HTGIGSVAFRLILKAHTLVKEIKGCTQIPQDSSASVSDTSQVYASKIEVDNEF				
gene	EDWQVATIEKVDQDFPNTVIGGSSNEMWGKDIPIATDVASDANGVAAGKKVCK				
CDS	MEDDEERTYATITETIKIQGNQVFIHTGYDSSIDENVNDDIS"				
gene	1680..1880				
CDS	/gene="asr4594"				
gene	1680..1880				
CDS	/gene="asr4594"				
gene	/note="ORF ID: asr4594				

```

hypothetical protein"
/codon_start=1
/translation="BAB76293.1"
/db_xref="GI:17133733"
/translation="MTAIAPEKSPROVLLYLGERDNYFVEVPSLPECISQKTRBEA
LTNIEVLQDKSETLPEYQFINI"
1877..2770
/gene="a1r4595"
1877..2770
/note="ORF_ID:a1r4595
unknown protein"
/codon_start=1
/translation="BAB76294.1"
/db_xref="GI:17133732"
/translation="MSKLVDRILLFQFTFVLGPTFTTHNVOASAPYMEFTNVDA
VQNGDMVITQCKNFRDYNHLSRYIPLDNDRTTEVSGENGLLSITSTNG
NHQLISWEHQLQSPESHFTVLKRVIGLLVNNHYQVWEALFSDRKAPIQAOVR
VEPELNTNKKIKIFQSPGVADIRVDTKITEAVTKQFIEPTELKQVIFDRAGTKI
ETSGMSTQPFYERLWILLGLVFGIPLLLIIVSRNGTQTSSTSIDSSIDYGGGG
DYGGDYGGDYGGGGGGGCD"
3082..4473
/gene="a1r4596"
3082..4473
/note="ORF_ID:a1r4596"
/translation="malic enzyme"
/product="malic enzyme"
/protein_id="BAB76295.1"
/db_xref="GI:17133733"
/translation="WADITENSFSVTLRIOIPNRVGMASITQALSSGCGQLD
IEQTVSLRDLTVDASTEHEFTIYQAVKAIPIKLVSYDRFTFNHARGKISITR
IPLKSVSDLAAYTVGVGVCTAIAODPSQVYNTLIQNTVAIVTDSGAVLGSGNFR
HAALFVMEGKAMLPKEPAGLDAPICLACTPTEIIQAVNKIAPVSGVMEIAPR
CEIEKRLRDELIDVPHDQGTALVTALFNSLKVQKSLADIRIVNGAGAVY
AIALIRLRAGAEEKIMQDSKGIISTRTDNEEKQEPAVKAGGLAGVAGAVFTIGV
SAPGILTPKMKSAKDAIVRANPPEIPELISQDVAVIANGRPDYNQINNTLA
FPGVFRGALDCAQTITMTLEASALASLVNPSDLRHHIIPSVDEVVTVAVAA
VQRAREGIARG"
5270..6853
/gene="a1r4597"
5270..6853
/note="ORF_ID:a1r4597"
/codon_start=1
/translation="type I restriction-modification system DNA
methylase"
/product="type I restriction-modification system DNA
methylase"
/protein_id="BAB76296.1"
/db_xref="GI:17133734"
/translation="VTEKAKQNNSESFEQKWKADKKNIDAAFYKHIVGLFTL
KYSDAPELHOKLLAGBDYAGANPEDRDYSANVFVPEARMSYLDGQAKQPI
KXTVDLAELAEQENAKRLKILPRVYGOQLDQKSGGLDLSITLIDAEQAD
VLGRVYEYFLQFALABGKGGQYTPESIVKLLVEMLEPNRVPDPCCGSGMFOQ
SEKFKVHQGRGLDISIYGQSNETTVKLCMNLAIKIGIDGSIKNPDESFLNDAK
DKLADFIANPNDSDWGEILLRNDKRLDQDLVPPGNANFAMVSHFYHLAPTS
AGVYLSNGSLSSNPSGEGDILKALVQDLYDCTMLPTCLFYNTGIPACIMPLSRYN
CKNRDRRGVEYFLIDASELGYMNARSSAPFRAEISKIDPTHEMKSGGSYVDICF
CKASIAIEKXNFVLTGRTVGI.PDEVEDGVSLEKMSLTMALGQWMEGQULDEE
IKQLLKVGFIQSDLVDF"
6858..7121
/gene="a1r4598"
6858..7121
/note="ORF_ID:a1r4598
unknown protein"
/codon_start=1
/translation="BAB76297.1"
/db_xref="GI:17133735"

```

```

gene
CDS
/translation="MNMTQVILKKNLPIVIEKLKHLAOSHORTLIEEITSILDEVTE
NTPITSKSRDWSPGFEQTCAGWQELLVREPOEADREPLL"
7226..7804
/gene="a1r4599"
7226..7804
/note="ORF_ID:a1r4599
hypothetical protein"
/codon_start=1
/translation="BAB76298.1"
/db_xref="GI:17133736"
/translation="MIAYKKPPLTPEEYFIEWEKQOLRHEYLNGEYVANTGTONH
GRASNITFIVKGHRRGGCCQVNSDCGINFERKDYVPPVSTQCDERDTALQAO
YFCLIVEVLSPTASYDRGDKRLFRNPSQDYLLVDAEKIALIDLKAKRGNWEIF
NYQLGDNIKLOSIDLSPFISBYEDIVVEESV"
7855..8145
/gene="a1r4600"
7855..8145
/note="ORF_ID:a1r4600
unknown protein"
/codon_start=1
/translation="BAB76299.1"
/db_xref="GI:17133737"
/translation="MIFKRNMTQVILKKNLPIVIEKLKHLAOSHORTLIEEITSIL
E"
8211..8573
/gene="a1r4601"
8211..8573
/note="ORF_ID:a1r4601
hypothetical protein"
/codon_start=1
/translation="BAB76300.1"
/db_xref="GI:17133738"
/translation="WKKLTQHOTATGYSVWHEHLLFGCLRLVESQKRRLLEYINQ
IPLKMTILNYDLKAOHMAQEBARLSKMGKXPAFIDQIASIAVSNLILVTNNVSDF
BSFNLAIVENYVNSGEG"
8577..9749
/gene="a1r4602"
8577..9749
/note="ORF_ID:a1r4602"
/codon_start=1
/product="type I site-specific deoxyribonuclease chain S"
/protein_id="BAB76301.1"
/db_xref="GI:17133739"
/translation="MSEWKEITLGEIADIINGOSPTEGTCNNNGQLPLNGPTEFGD
RNPILPTQFTIDPKLIAAGADLFCVRSSTGRMNAADQVAIGRGIASIRADGILFQ
PYIRAIIEKEKSLIAVANGSTPMISDHLNLIYVLPNSNIKIYISNLARIDBKI
YNRKQNTLEIAIOTLFRHMTIDPEPNADGKPKYSGSGAMVSAAGYIPAMSVK
IAQVYINIKHGIKGEYITTEYTEKILITPPNFKIIGSFNDSKTYISADYSNYTL
RKRDIAITDTSKRGDSLGPAPFIDIKGVLPFHNOIGKVENNNIKITLIFULCR
RBYRHIIGTSSGSTRHSPRICEYFVIDPELIDKFSALATATLIDKILLISTL"
9944..10465
/gene="a1r4603"
9944..10465
/note="ORF_ID:a1r4603
unknown protein"
/codon_start=1
/translation="BAB76302.1"
/db_xref="GI:17133740"
/translation="WIGKRLNDRNRSFHHIALAESQAEKYLILSKYIDILFEV
KLESISLQIKNTLSNITDTRFLKFDHSLKKRHHDIYDNNPFRITNDINLEIRLF

```

Query Match 48.2%; Score 272.2; DB 1; Length 342850;
Best Local Similarity 71.9%; Pred. No. 1,le-65;

Matches	383;	Conservative	0;	Mismatches	148;	Indels	2;	Gaps	2;
OY	33	GCCTCTTTTCC	CCCCATGAAACACACACCCCTCTCTGTTTATGTTGAAGATGAAGCCGGAGT	92					
Db	39898	GCCCTGTGTGGCAAAAGTGAACATACTCTTTCGTGTGTGTGAAGAGTGAAGCCGGAGT	39957						
OY	93	GCTAACCCCGATTGCCGACTATTTTGGCCCGCTGGTTTTTAACATTGAGCTTGCGCGT	152						
Db	39958	TCTGTCCCGATTGTCTATGTTATTTGTGTGTGTGTCTTAAATATGAAGCCCTGCTGT	40017						
OY	153	GGGCTCGGCGGACAGGGGAGCGTTTCCCGCATCACACATGCTGTGCCCGGGAGTGAAGA	212						
Db	40018	TGTTTCGGGGGAGAGGGCGGAGTTTCCCGCATTAATGATGTTGATCTGGTGAAGATCG	40077						
OY	213	CACATGAAACAATGACACAGCAACTCTCAAGTTGGTTAACGTATTAATAGTACAGGA	272						
Db	40078	CGTATTTGAACAATCACCAAAACAATATCAAGCTCGTTAACTCTCAATGATCAGGA	40137						
OY	273	CATCACGGAACCTCCCTGTGTGGAAGGGAATTGATGCTGTGGAAGTGAAGGCCCAATGC	332						
Db	40138	TATACGTAACTCTTGTGTGTGAAGAGATGATGCTTCTGAAAGTGAAGCCACGAG	40197						
OY	333	CCCTAACCGAGCGGAATGATGATGAGCCGACGATATCCGGGCCCGCATTTGATAT	392						
Db	40198	CAGAACCGTTCAAGATCTATGATTAAGTCAAAATTTTCCGGGGCGAGTGTGATGT	40257						
OY	393	CTCCGAAGACACGCTCACCATCGAA-TGTGTGGGGAGCCGGGTAAATGTGTGCAATCT	451						
Db	40258	GGCGGAAGATCTTTTAACTTTAAGAGTGCGCGACCTGTGTAAGATGTGCGCATGCT	40317						
OY	452	CCAGATGTGGCCAG-TTGGCATTAAAGAGTGCGCTCGACGGCGAAATGCTTTGCT	510						
Db	40318	CCAGTGCTACAAAGTTTGTGTATCAGAGAAATTTGCCCGCACCGGTAAATGCCCTGAC	40377						
OY	511	GCGGGAATCCGGCGTCAATACGGAATATCTGAAATCCCTTGAATCCAACTTT	563						
Db	40378	TGGTAATCTGGTGCTCAATAGTGAATTAATCTTAAGAGCGAAGTTT	40430						
RESULT 3									
AP005371/c	AP005371	300450 bp	DNA	linear	BCT 17-AUG-2002				
LOCUS	Thermosynechococcus elongatus BP-1	complete genome, section							
DEFINITION	3/9.								
ACCESSION	AP005371	BA000039							
VERSION	AP005371.1	GI:22294315							
KEYWORDS									
SOURCE									
ORGANISM	Thermosynechococcus elongatus BP-1								
REFERENCE	Thermosynechococcus elongatus BP-1								
AUTHORS	Bacteria; Cyanobacteria; Chroococcales; Thermosynechococcus.								
	1 Nakamura, Y., Kaneko, T., Sato, S., Ikeuchi, M., Katoh, H., Sasamoto, S.,								
	Watanabe, A., Iriyuhui, M., Kawashima, K., Kimura, T., Kishida, Y.,								
	Kiyoekwa, C., Kohara, M., Matsuno, M., Matsuno, A., Nakazaki, N.,								
	Shimpo, S., Sugimoto, M., Takeuchi, C., Yamada, M. and Tabata, S.								
TITLE	Complete genome structure of the thermophilic cyanobacterium								
JOURNAL	Thermosynechococcus elongatus BP-1								
REFERENCE	DNA Res. (2002) In press								
AUTHORS	2 (bases 1 to 300450)								
TITLE	Kaneko, T.								
JOURNAL	Direct Submission								
	Submitted (05-JUN-2002) Takakazu Kaneko, Kazusa DNA Research								
	Institute, The First Laboratory for Plant Gene Research; 2-6-7								
	Kanusa-kamatari, Kisarazu, Chiba 292-0812, Japan								
	(E-mail:kaneko@kazusa.or.jp,								
	URL:htp://www.kazusa.or.jp/cyano/thermo/								
	Tel:81-438-52-3935(ex.2338), Fax:81-438-52-3934)								
FEATURES									
source	1. 300450								
	/organism="Thermosynechococcus elongatus BP-1"								
	/mol_type="Genomic DNA"								
	/strain="BP-1"								
	/db_xref="taxon:197231"								

gene
126..1655
/gene="t1r0594"
126..1655
/gene="t1r0594"
/note="ORF_ID:t1r0594
competence proteincm Com homolog"
/codon_start=1
/transl_table=1
/protein_id="BAC08146.1"
/db_xref="GI:22294316"
/translation="MKLRVMSAAVVGIDAFVGVYVYVSGGLPRTYVYVGLPDAGVQAR
KERVYDAIRNNGSPFMGQILVNLTLPADLRKEGSPFLPISIGILAAAGVATDILGDL
HLFDELSIDDTLOCPVAGVLAIALAAQAQGITGLVVPATNATEALVGLKRVYCHTL
AEVAPGSKTDPSSSPATPSPCMVAAPAEFVETKHSVAGGLKERGALEIAAGGHNLI
FVGPAGSFMDIMARLPITLPPLFAAPAEVTKHSDVAGGLKERGALEIAAGGHNLI
SASGALVAGGSGYPRPGEISLARGVFLDELTFEKKDVEILEQPLEDCQVSIARAR
OSVAPPAFTIVASTNPGCGYVGDVPVQCTCSPROQYQWAKLSGPLRLIDQVSV
SRKKEBMTROPPLGSDSTYRQVILARSPAQCPAAEPPVNHCAQMOASHIRPMGSL
DDASHLHERAIQDLSARATDRILKYARTILALACETTTSNHAVALQYRIIDRL
Q"

gene
1588..2898
/gene="t1r0595"
1588..2898
/gene="t1r0595"
/note="ORF_ID:t1r0595
probable sugar transporter sugar binding protein"
/codon_start=1
/transl_table=1
/protein_id="BAC08147.1"
/db_xref="GI:22294317"
/translation="MKQPLPLPMWRKRPENTAOITVCNRRGWFALLLGLFATLVILICTTA
TQSEQPIPLMHGVIQVPPENRVYLORLVDRENARHPQIHVOALVVGPDQDLPKILAA
VYGNAPDLIMVHGVNPGITGQPFELGALRPIDDMWQASPRDQVSPALIPIWRVGNHWS
IPFANNIGITYRRPSLPPAAAIKTLIPSTWQLEVOYSQITRAQGITPLILPGQSEFAY
FMSLFFPSAGSGSDEIETETAHIDTPAMALBERWRLKRGGLIISAPERGITLDQF
RGEVAAQMTIDPWTGLQEQSGVDVFGVLPWMLTPTPAALQGENLFFVRSIPQKEBA
LAFILVYVSEBQETMALEGTGVLPIINERVINDYRPAFVAQGNPLRVFLQMAAAQAR
PNFRGAFRSOTLGLFALESVLLGKSTPKCAVTTAEORWQLMR"

gene
complement(2895..3905)
/gene="t110596"
complement(2895..3905)
/gene="t110596"
/note="ORF_ID:t110596
probable heme di biosynthesis protein NirJ"
/codon_start=1
/transl_table=1
/protein_id="BAC08148.1"
/db_xref="GI:22294318"
/translation="MGVTRKAVENVATYIISGRLOKKRPPLTLMLEELPRONLACSG
GCKIHLELVAKKTLTPEECFRVAVBEGCAPVVALPGEPLHLPQIDELVSLVARRKK
VYLCTNGLLIBENLKRKPSYFSFSVHLDELIRMHDCVARKGVFTDAVNAIQAAR
KQFRVTYITLTFEGVDPRKIOEFDFVSLSDVDMMSIPGAYWADQDFHFLREOTD
RALFEELIAPYAGQKKNPNHNPFLFDLFLGEXDYECTPMGMSYSYVLQKQKCYDIL
NGCHYKTWQELIENTMDIKYGHKSGNPKCRDQCMYHCGYEATAMDAEPANMVAAASS
LF"

gene
4441..6447
/gene="tne"
4441..6447
/gene="tne"
/note="ORF_ID:t1r0597"
/codon_start=1
/transl_table=1
/product="ribonuclease E"
/protein_id="BAC08149.1"
/db_xref="GI:22294319"
/translation="MPKQITIAEHRRLAAVPAEDIOQLVATGTHOVGDIYLGAVENR
VLPGLDAFVNIGDSNRNGFIVHSDLGFLRLKATGATITELLAPOQVLAQVWKEPFG
TGPGLTANLSPGRYIVLMPYGRGVNLSRRTTEAEHRRLALGILVLPKMGELVRA
TAEVSESAILEDLELQRMETIQCCAASRPOLGRDDRIQKRLVYSSDVAV
RLAVTPDGVKQKHOHLKMMVNNRPVGLIDYHQBPIILDYFVNALIRALKRPVYD
LBSGGVITIOPEALFTVDVNGSFTSATRETLVWNCGAETELROULENLAIRP
ITVDITNDSPDQDITTFHFSKATRAKARAPDIAQISGILVFETRQOCNTYPLE

gene
4441..6447
/gene="tne"
4441..6447
/gene="tne"
/note="ORF_ID:t1r0597"
/codon_start=1
/transl_table=1
/product="ribonuclease E"
/protein_id="BAC08149.1"
/db_xref="GI:22294319"
/translation="MPKQITIAEHRRLAAVPAEDIOQLVATGTHOVGDIYLGAVENR
VLPGLDAFVNIGDSNRNGFIVHSDLGFLRLKATGATITELLAPOQVLAQVWKEPFG
TGPGLTANLSPGRYIVLMPYGRGVNLSRRTTEAEHRRLALGILVLPKMGELVRA
TAEVSESAILEDLELQRMETIQCCAASRPOLGRDDRIQKRLVYSSDVAV
RLAVTPDGVKQKHOHLKMMVNNRPVGLIDYHQBPIILDYFVNALIRALKRPVYD
LBSGGVITIOPEALFTVDVNGSFTSATRETLVWNCGAETELROULENLAIRP
ITVDITNDSPDQDITTFHFSKATRAKARAPDIAQISGILVFETRQOCNTYPLE

VERSION	AP006572.1	GI:35211683
KEYWORDS	Gloeobacter violaceus PCC 7421	
SOURCE	Gloeobacter violaceus PCC 7421	
ORGANISM	Bacteria; Cyanobacteria; Chroococcales; Gloeobacter.	
REFERENCE	1 Nakamura, Y., Kaneko, T., Sato, S., Mimuro, M., Miyashita, H., Tsuchiya, T., Sasamoto, S., Watanabe, A., Kawashima, K., Kishida, Y., Kiyokawa, C., Kohara, M., Matsumoto, M., Matsuno, A., Nakazaki, N., Shimpō, S., Takeuchi, C., Yamada, M. and Tabata, S. Complete genome structure of Gloeobacter violaceus PCC 7421. a cyanobacterium that lacks thylakoids (supplement)	
TITLE	2	
JOURNAL	DNA Res. 10, 181-201 (2003)	
REFERENCE	2 Nakamura, Y., Kaneko, T., Sato, S., Mimuro, M., Miyashita, H., Tsuchiya, T., Sasamoto, S., Watanabe, A., Kawashima, K., Kishida, Y., Kiyokawa, C., Kohara, M., Matsumoto, M., Matsuno, A., Nakazaki, N., Shimpō, S., Takeuchi, C., Yamada, M. and Tabata, S. Complete genome structure of Gloeobacter violaceus PCC 7421. a cyanobacterium that lacks thylakoids	
AUTHORS	DNA Res. 10, 137-145 (2003)	
TITLE	3 (bases 1 to 299350)	
JOURNAL	Submitted (15-Aug-2003) Takakazu Kaneko, Kazusa DNA Research Institute, The First Laboratory for Plant Gene Research; 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba 292-0818, Japan (E-mail: kaneko@kazusa.or.jp, URL: http://www.kazusa.or.jp/cyano/, Tel:81-438-52-3935(ex.2338), Fax:81-438-52-3934)	
REFERENCE	Location/Qualifiers	
AUTHORS	1. 299350	
JOURNAL	/organism="Gloeobacter violaceus PCC 7421"	
TITLE	/mol_type="genomic DNA"	
REFERENCE	/strain="PCC 7421"	
AUTHORS	/db_xref="taxon:2511221"	
JOURNAL	complement(1.. 1224)	
TITLE	/gene="g111120"	
REFERENCE	complement(1.. 1224)	
AUTHORS	/gene="g111120"	
JOURNAL	complement(1.. 1224)	
TITLE	/note="ORF_ID:g111120"	
REFERENCE	unknown protein"	
AUTHORS	/codon_start=1	
JOURNAL	/protein_id="BAC89061.1"	
TITLE	/db_xref="GI:35211684"	
REFERENCE	/translation="MKLTGVNPKYKGRFMRNREKELNGLIQLDDEKAHLLIAQRICK TSLNEQABITIKRILYLCQIDQSAHVSVDIAETLAETLRSLSMEKLTGTFNSVLN AARQIDLSQLSELKVTLLFTGMTSGMKAQGLPKLAESDKPVIPLDEVPILIVSR LKSGDYMTPEPRGQTDALFSLWRANSJLKHREKIRLLVVTGSLGEPILRQVLSATIN HLTSPISIGMEPTVVKSCVEALAEYSLTFGDGLGRCVPIVQQLPFOQVRA TVCVAREGLTHIKPILVAEYEREMLAGRGAELSHLEERIKLIVIGPELHPALFFLES AAVTGKLSAALGVCFLEVKQPVVIARILLILIEHDGIMYQOSNTYRRVSKLVGBWR AOYSLFVPAASRGK"	
AUTHORS	1511.. 1966	
JOURNAL	/gene="g1r1121"	
TITLE	1511.. 1966	
REFERENCE	/gene="g1r1121"	
AUTHORS	/note="ORF_ID:g1r1121"	
JOURNAL	hypothetical protein"	
TITLE	/codon_start=1	
REFERENCE	/transl_table=11	
AUTHORS	/protein_id="BAC89062.1"	
JOURNAL	/db_xref="GI:35211685"	
TITLE	/translation="MAEYDASALERAIESAREAGAROLAGPLAERVYDGRSRLTVI YLNTGVMIPVRSLOAGATSOOLREVTGSPRTGIRFPALDAVFLDGLNGIFG TERNMMDLRTAAKGSARSAKINAAANAGKGRPKQKQPHLVVAPE"	
REFERENCE	2128.. 3642	
AUTHORS	/gene="g1r1122"	
JOURNAL	/gene="g1r1122"	
TITLE	2128.. 3642	
REFERENCE	/note="ORF_ID:g1r1122"	
AUTHORS	probable long chain fatty acid CoA ligase"	
JOURNAL	/codon_start=1	
TITLE	/transl_table=11	
REFERENCE	/protein_id="BAC89063.1"	
AUTHORS	/db_xref="GI:35211686"	
JOURNAL	/translation="NMLPLVHRAYEHGPRATVADGCFSYODLLSSQVATRLDG ADPLREQVRASLVPAGREYVATQMGIMRAGGIAVPLCVSHPRRELEYITNSGASIV IAPHSFSLRPLIAEHAQRLPILTSBSLPPTDQLPVEPTRALITITSGTTRKRGV VTHNGIQAQVTSLSISAWETSQDRILHVLPLHILHIVNLICAMAGABCHMSKFL DAEIVHEHICEGFTLFMAVPTIYTKLAIHWKAAQDRQMDSRGCAKRWLVSSGSA LPPQVLEWTHLSGHFLERYGTETIGMALNSPLHGRRAGYGMALPQDVVALDES GSVDPAGTGEIGVKGQGVFLRWQKPEATAKAFODGFCGTGDLAVEDDCYALIGRI SVDIITKGKVASLAEIEEVLRTHPDIOCAVGVAGDPEMGSEVCGALVLOGCNLAL EPRRSAXERLAAVKYPTRTILSYEBELPRNAMGKVTKEPAVAQLPRAQIABSDPA"	
REFERENCE	complement(3775.. 5196)	
AUTHORS	/gene="g111123"	
JOURNAL	complement(3775.. 5196)	
TITLE	/gene="g111123"	
REFERENCE	/note="ORF_ID:g111123"	
AUTHORS	unknown protein"	
JOURNAL	/codon_start=1	
TITLE	/transl_table=11	
REFERENCE	/protein_id="BAC89064.1"	
AUTHORS	/db_xref="GI:35211687"	
JOURNAL	/translation="MPSIAITGGLAGLQAAVYOLKKGUYADVYEAKSFVGGRIQSRK NVVPGIAYVEIGGTFINSHRDLITLTAQSENLPLFRADAQOFPPEPTGYFDDGNR LSEVAREKURPLADQIAKDAALIDENFDKFAPOQFRLSVARVYINLHAKTIPEPTITL IENSIITREYGEPEYDSSAQLQPLNLEPVNDGEAVEYSGSDEITFVVGSGERLIASIAA VHGRIHTNRALTRKMGQKRLTFAGAQTRVEDVTIVAVPNGLKKTIDFOVPLDPA FRFRIEVSIGFNEKLPAGSERVWMOEESGVBGWTJGFSQWEDTQRQPSDAA LTFPFMAQOEFAFQSLGAQTLGEAPLQFEAAVVDGQAATGFRSXTNNQDDPWIGG YTFPFKQGLTEFGFPLYIESVILPAFRQDVNFNLIPAGHLSDEFFYFMNAAQTREL AAOVVLROEOEGGLASKPTSA"	
REFERENCE	5942.. 6382	
AUTHORS	/gene="g1r1124"	
JOURNAL	5942.. 6382	
TITLE	/gene="g1r1124"	
REFERENCE	/note="ORF_ID:g1r1124"	
AUTHORS	hypothetical protein"	
JOURNAL	/codon_start=1	
TITLE	/transl_table=11	
REFERENCE	/protein_id="BAC89065.1"	
AUTHORS	/db_xref="GI:35211688"	
JOURNAL	/translation="MLSDVDVVRIVVDTSVFISALIGAPGSRRLVRRCLOGQYPL MGNALPCEYSIVGRPHIILACPLSQNEILITLHMLMSYKMSIYYLWPRMLKDBAD NHIIELVAGNAQVIAITNNVRDPRONALLFPDLSILKPEIITIR"	
REFERENCE	6389.. 6607	
AUTHORS	/gene="g1r1125"	
JOURNAL	6389.. 6607	
TITLE	/gene="g1r1125"	
REFERENCE	/note="ORF_ID:g1r1125"	
AUTHORS	hypothetical protein"	
JOURNAL	/codon_start=1	
TITLE	/transl_table=11	
REFERENCE	/protein_id="BAC89066.1"	
AUTHORS	/db_xref="GI:35211689"	
JOURNAL	/translation="VATITRLPDDYHERIKELAQTRGISVNKLIELSTIALAEFDA YTRFKMAAAGNPESGLKLDKDALTG"	
REFERENCE	complement(6672.. 7328)	
AUTHORS	/gene="g111126"	
JOURNAL	complement(6672.. 7328)	
TITLE	/gene="g111126"	
REFERENCE	/note="ORF_ID:g111126"	
AUTHORS	hypothetical protein"	
JOURNAL	/codon_start=1	
TITLE	/transl_table=11	
REFERENCE	/protein_id="BAC89067.1"	
AUTHORS	/db_xref="GI:35211690"	
JOURNAL	/translation="MRKGRANOTPFVODAMWRASMESRTALTNDPAIVHAKMYHL GWRIRKSPVGPAAEDNNLIAOVVQIFAFSRGIRLKGTYINLTIRKTELQEAQPDLAY YLENOVSPRPARINSFVLDLQIAAPALAEISATLDDQKRELYAQIGVGGEYVAVD AENTQVLLPGPIPGVDSLEPIVASVLPGLTATLCEALRLQSGESGDAAAIRVDLE "	
REFERENCE	complement(7361.. 8260)	
AUTHORS	/gene="thr8"	

CDS	complement(7361..8260)																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																	</
-----	------------------------	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	----

Query Match 41.7%; Score 235.6; DB 1; Length 299350;
Best Local Similarity 67.7%; Pred. No. 2.6e-55;
Matches 359; Conservative 0; Mismatches 169; Indels 2; Gaps 2;

QY	23	ACCGCGCATGCCCTTCTTGGCCCCCATGAAACACACCCCTCTGCTTTTGTGAAGATG 82
DB	15531	ACTATGGCGCATTTGAGAGCGAGCGTAAGCATACCTTGGGTATTAAGTGAAGATG 15472
QY	83	AAGCGGAGTGTCTAACCCGATTCGCGACTATTGGCCCCCGTGTGTTTAACTATGAGA 142
DB	15471	AGGCGGCGGCGCTACCGCGATGCGCGCATGTTCCGCCGCGGCTTCAATATGAGA 15412
QY	143	GCTTGGCGGTGGGCTGCGCGGACAGGGGGAAGCTTCCCGCATCAACATGCTGCTCCG 202
DB	15411	GCTTGGCGGTGGGCTTACCGGAGCAATCGGCGACACCCCCATCAAGATGCTGCTGCGG 15352
QY	203	GGATGAGAACACCATGAGAACACTGACCAAGCAACTCTAAGTGGTGTAACTATTA 262
DB	15351	GGGACGAAACAGGCGGTGAGCAGATTCACAAAGCAGCTGCAAGCTCATACAGTCTGCTC 15292
QY	263	AAGTACGAGCAATCAACGAAACCTCCTGTGTGAAAGGAATTGATGCTGTGAAGTGA 322

DB	15291	GTGTGAGAGCAATCAACCAATTCCTCTGCTGAGCGGGAACGTGATCATCAAGATGA 15232
QY	323	GGCCAAATGCCCTTCAACGAGCGGAGATGATGAGTCAAGTATTCGGGCGCCCA 382
DB	15231	AGCCACACCGCCCGACCGGACGAAATCATGAGTATGCGGCGCCCA 15172
QY	383	TGTGATATCTTCGGAAGACACCGTCAACCATGA-ATGTTGGGAGACCGGCTAAATG 441
DB	15171	TGCTGATGTGGCGGAAGATGTCACCGCTGAGATGACCGGCATCCGGGCAAGATG 15112
QY	442	TGCAATCTCTCAGATGTTGGCCAAAGT-GGCATTAAGAGAGTGGCTGAAACGGCAAA 500
DB	15111	TGCGCATATTCAGATGCTCAACCGCTTCGGCATCAAGAAATTCGCCGACCGCAAAA 15052
QY	501	TTCCTTGGTGGCGGAATTCGGCTCAATACGGAATATCTGAATCCCTG 550
DB	15051	TTCCCTACACCGCGAGTGGGCTGAACACCGAATTTCTGCGCACCCAG 15002
RESULT 5		
BX569690/c		
LOCUS	BX569690	349652 bp DNA circular BCT 23-SEP-2003
DEFINITION	Synechococcus sp. WH8102 complete genome; segment 2/7.	
ACCESSION	BX569690 BX548020	
VERSION	BX569690.1 GI:33638574	
KEYWORDS	complete genomes.	
SOURCE	Synechococcus sp. WH 8102	
ORGANISM	Synechococcus sp. WH 8102	
REFERENCE	Bacteria; Cyanobacteria; Chroococcales; Synechococcus.	
AUTHORS	1 (bases 1 to 349652) Palenik,B., Brahmasha,B., Larimer,F., Land,M., Hauser,L., Chain,P., Lamerdin,J., Regala,W., Allen,E.A., McCarren,J., Paulsen,I., Dufresne,A., Poretsky,F., Webb,E. and Waterbury,J.	
TITLE	The genome of a motile marine Synechococcus	
JOURNAL	Nature 424 (6932), 1037-1042 (2003)	
MEDLINE	22825697	
PUBMED	12917641	
REFERENCE	2 (bases 1 to 349652) Larimer,F. and Palenik,B.	
AUTHORS	Synechococcus genome consortium	
CONSTRM	Direct Submission	
TITLE	Submitted (01-JUL-2003) Larimer, F., Submitted on behalf of the Synchococcus genome consortium, the DOE Joint Genome Institute, Production Genomics Facility, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA, and the Genome Analysis Group, Oak Ridge National Laboratory, 1060 Commerce Park Drive, Oak Ridge, TN 37831, USA; larimerf@ornl.gov	
FEATURES		
source		
1..349652		
/organism="Synechococcus sp. WH 8102"		
/mol_type="genomic DNA"		
/strain="WH8102"		
/db_xref="taxon:84588"		
complement(173..1291)		
/locus_tag="SYNM0355"		
complement(1173..1291)		
/locus_tag="SYNM0355"		
/codon_start=1		
/evidence=not experimental		
/transl_table=1		
/product="hypothetical"		
/protein_id="CAE06870.1"		
/db_xref="GI:33638575"		
/translation="MITTALPRMGSGRISEAKLQPHADNEHVVAMRQYVARIGPAP VARSWLYALBESGIIITKGDFTWGSXWADRRKGLPFELVGLADSTLALSGSDVYDKE ETPREYINRQLEBSLIDOKHWPASWYKHQYEPIIWERKDLIKLEPELPAIRRF ASGQADVNSRVALIOECDMALQGLPEPVLYGDLDPMVYKNSDALAEHLRPLAKYL NMDSLIDMDMDGRIIVRFGNAEFDVNAIGLVLDGLETSSGDLANPNHHRKRPVYO DYLETYGRKREANALANPRAARSILRDELWMLSHSGKOWSENRKASSEATGIV DAITRMALMFDSVGLATNPRLEQAVNSGLNSLPAA"		
complement(1388..3358)		
/locus_tag="SYNM0356"		

gene

Query Match	Similarity	Score	DB	Length	8200;
Best Local	65.0%;	Pred. No. 6e-44;			
Matches	319;	Conservative	0;	Mismatches	170;
				Indels	2;
				Gaps	2;
ORIGIN					
Query Match	34.5%;	Score 195;	DB 8;	Length 8200;	
Best Local	65.0%;	Pred. No. 6e-44;			
Matches	319;	Conservative	0;	Mismatches	170;
				Indels	2;
				Gaps	2;
ORIGIN					
Query Match	34.5%;	Score 195;	DB 8;	Length 8200;	
Best Local	65.0%;	Pred. No. 6e-44;			
Matches	319;	Conservative	0;	Mismatches	170;
				Indels	2;
				Gaps	2;
ORIGIN					
Query Match	34.5%;	Score 195;	DB 8;	Length 8200;	
Best Local	65.0%;	Pred. No. 6e-44;			
Matches	319;	Conservative	0;	Mismatches	170;
				Indels	2;
				Gaps	2;
ORIGIN					
Query Match	34.5%;	Score 195;	DB 8;	Length 8200;	
Best Local	65.0%;	Pred. No. 6e-44;			
Matches	319;	Conservative	0;	Mismatches	170;
				Indels	2;
				Gaps	2;
ORIGIN					
Query Match	34.5%;	Score 195;	DB 8;	Length 8200;	
Best Local	65.0%;	Pred. No. 6e-44;			
Matches	319;	Conservative	0;	Mismatches	170;
				Indels	2;
				Gaps	2;
ORIGIN					
Query Match	34.5%;	Score 195;	DB 8;	Length 8200;	
Best Local	65.0%;	Pred. No. 6e-44;			
Matches	319;	Conservative	0;	Mismatches	170;
				Indels	2;
				Gaps	2;
ORIGIN					
Query Match	34.5%;	Score 195;	DB 8;	Length 8200;	
Best Local	65.0%;	Pred. No. 6e-44;			
Matches	319;	Conservative	0;	Mismatches	170;
				Indels	2;
				Gaps	2;
ORIGIN					
Query Match	34.5%;	Score 195;	DB 8;	Length 8200;	
Best Local	65.0%;	Pred. No. 6e-44;			
Matches	319;	Conservative	0;	Mismatches	170;
				Indels	2;
				Gaps	2;
ORIGIN					
Query Match	34.5%;	Score 195;	DB 8;	Length 8200;	
Best Local	65.0%;	Pred. No. 6e-44;			
Matches	319;	Conservative	0;	Mismatches	170;
				Indels	2;
				Gaps	2;
ORIGIN					
Query Match	34.5%;	Score 195;	DB 8;	Length 8200;	
Best Local	65.0%;	Pred. No. 6e-44;			
Matches	319;	Conservative	0;	Mismatches	170;
				Indels	2;
				Gaps	2;
ORIGIN					
Query Match	34.5%;	Score 195;	DB 8;	Length 8200;	
Best Local	65.0%;	Pred. No. 6e-44;			
Matches	319;	Conservative	0;	Mismatches	170;
				Indels	2;
				Gaps	2;
ORIGIN					
Query Match	34.5%;	Score 195;	DB 8;	Length 8200;	
Best Local	65.0%;	Pred. No. 6e-44;			
Matches	319;	Conservative	0;	Mismatches	170;
				Indels	2;
				Gaps	2;
ORIGIN					
Query Match	34.5%;	Score 195;	DB 8;	Length 8200;	
Best Local	65.0%;	Pred. No. 6e-44;			
Matches	319;	Conservative	0;	Mismatches	170;
				Indels	2;

QY	288	CTGTGTGAGAAAGGGAATTTGATGCTGTGAAGGTGAGCGCCAAATGCCCTTAACCGGACGGA	347
QY	7847	TTGTGTGAGAAAGGGAATTTAATTTTATAGAGGTGGAACCAATTCCTCCAGACGAGGAGAGA	7788
QY	348	AGTGAATTAGCTAGTCCCGGAGTATTCGGGACCGCCGACCTGTGGATATCTCCGAGACACCGGT	407
Db	7787	AATTTTAATATATGCGTAAATATATTTTTCGGGACAGTGTGTGTGATATATATACGACGAGCATGT	7728
QY	408	CACCATCGAA-TGGTGTGGGAGACCCCGGTATAATTGTGTAGCAATCTCCAGATGTTGGC-CA	465
Db	7727	CATGTATAAGATTAACGGGTGATCTCTGGAATAATGTTAGCATCTGAAACAAATTTTAAAGAA	7666
QY	466	AGTTGGCATTTAAAGAGTGGCTCTCGAACGCGCAAAATTCCTTGTGTGGGGAATTCGGCGGT	525
Db	7667	ATTGTGTATAGTGAAGTGAAGTACCGCGTACGAGAAATAATTCATTTAAACGATCATCGAAGT	7608
QY	526	CAATACGGAT	536
Db	7607	TAAACGAGAT	7597
RESULT 8			
AB002583/c		149987 bp	DNA
LOCUS			circular PLN 07-MAY-2003
DEFINITION			Cyanidioschyzon merolae chloroplast DNA, complete genome.
ACCESSION			AB002583
VERSION			AB002583.1
KEYWORDS			GI:30409147
SOURCE			chloroplast Cyanidioschyzon merolae
ORGANISM			Cyanidioschyzon merolae
REFERENCE			Eukaryota; Rhodophyta; Bangiophyceae; Cyanidioschyzon.
AUTHORS			1. Ohta,N., Matuzaki,M., Misumi,O., Miyagishima,S., Nozaki,H., Tanaka,K., Shin-I,T., Kohara,Y. and Kuroiwa,T.
TITLE			Complete Sequence and Analysis of the Plastid Genome of the Unicellular Red Alga Cyanidioschyzon merolae
JOURNAL			DNA Res. 10, 67-77 (2003)
REFERENCE			2 (bases 1 to 149987)
AUTHORS			Ohta,N., Matuzaki,M., Tanaka,K., Shin-I,T., Kohara,Y. and Kuroiwa,T.
TITLE			Direct Submission
JOURNAL			Submitted (02-APR-1997) Niji Ohta, Saitama University, Department of Molecular Biology and Biochemistry, 255 Shimo-ohkubo, Saitama, Saitama 338-8570, Japan (E-mail:nijioh@molbiol.saitama-u.ac.jp, Tel:81-48-858-3848, Fax:81-48-858-3384)
FEATURES			Location/Qualifiers
SOURCE			1..149987
			/organism="Cyanidioschyzon merolae"
			/organelle="plastid;chloroplast"
			/mol_type="genomic DNA"
			/strain="10D"
			/db_xref="taxon:45157"
			1.._1398
			/gene="ycf24"
			1.._1396
			/gene="ycf24"
			/codon_start=1
			/transl_table=1
			/product="ABC transporter for iron-sulfur cluster formation (subB)"
			/protein_id="BAC76097.1"
			/db_xref="GI:30409148"
			/translation="MNQQLNDRMINREYAYGFTSDYEPPIGLNEQLVKMISAKKCEPQFMDFRLKAVHWKQMRPQWARYATINFEQLYYAASPTEDDVOSTLAKIGVKSFKALDPVDSVSLGTGFQALNDAGVITCSLEAIQKPSLVKXLAIVPDI GDNFSKALNVAISVDSGECYIPDTACMEASTYFRLKESGGFRLIIAEGASVSYLSCGATARYDRHLHAALVWKAALDHAISIKSTVOWYSGDQEGKGIINPYTKG LCAEGKATSWTQVETGSAITWELKALPCSLANNRADYVSLTDDQADPTGSKMHL IGANSKRITIKSIGTIGKSGSMNTYGGQVMTGAQAQAMHSGCDSEILCKKATNTPTPI IELANVSARVEHATITDITGEEQLFTLQQRGINOVADANQWVLGCGQVFLQEPNERPA TVEVELITKLENGC"

gene 1399. .2139
/gene="ycf16"
CDS 1399. .2139
/gene="ycf16"
/codon_start=1
/transl_table=11
/product="ATP-binding subunit of ABC transporter for
iron-sulfur cluster formation (suc)"
/protein_id="BAC76098.1"
/db_xref="gi:30409149"
/translation="MLKIEQLKVGKDLLEGLDAHIEAGSVAHIMQNGSGSTILA
KYIAHPYKLTIEGFKMDSINDKDAEAAAGIPLAFGNPLTIGVTMDMLLA
YTKHSGSDIASPLEYMPKLEIVGDASPIHAYVNDGSGKKRKEITOLAILEE
KLAIDEDISGDIDALQIAHSLMKLTQNAVYLITHPRLEIYIKPDVYHWOKK
RILISGDSLAKYERIVGDEKKAAY"
complement(2171. .2275)
/gene="ycf12"
CDS complement(2171. .2275)
/gene="ycf12"
/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="BAC76099.1"
/db_xref="gi:30409150"
/translation="MMHVLIAQLTVLAFVAVGPITLIMLNKDALL"
complement(2193. .3065)
/gene="ycf62"
complement(2193. .3065)
/gene="ycf62"
/codon_start=1
/transl_table=11
/product="cell-cycle protein meso family"
/protein_id="BAC76100.1"
/db_xref="gi:30409151"
/translation="YTPMHNKFNORPOLANLTSHSPSHSLAISGGDSQILTKRK
MHDIHHHLVAVHPHMSGSLAARVYLSRWHLAMRYFRTALPTTEDDAQVRY
ANLIDLYQVTHSHACTATKATSDLETLTLDWMISLRHEGWRCHSGSILIRPLL
HLTRQDFTLIVHTLPTWSDVSNQITFRKNQLRHQILPTKLNPEPFTTSLRPS
AHLSPLVPEHWVSSLPWLQVIAOTWGLTISOQLIYNKXELARSSINRRS
FQDSRSRYHFNAL"
3086. .3157
gene 3086. .3157
/gene="crnk"
CDS 3086. .3157
/gene="crnk"
/product="crnk-Lys"
/note="codon recognized: AAA"
/anticodon=(pos:3118. .3120,aa:Lys)
3183. .3914
gene 3183. .3914
/gene="crpA"
CDS 3183. .3914
/gene="crpA"
/codon_start=1
/transl_table=11
/product="tryptophan synthase alpha chain"
/protein_id="BAC76101.1"
/db_xref="gi:30409152"
/translation="MLIAYLTAGAPDINTTKEAVMKLAKGADVIEIGVPSYDLADQG
ALIQASQALMNGGKHLDMNLSEVVEIEVPVILAYYQIWHYGEVKKVCKLVAAH
NVKGIIVDDLPIRESKTLROI CDRYGLNITIMLISPTNKTAEGLARACKWIVIR
TGVLTLEHPKQIPIKIGELKVTXAPINLGFQISSEQYKLVKMSGADVIIIGSAC
MOLLERKVDQLSEWISWKSFCP"
3800. .5140
gene 3800. .5140
/gene="infB"
CDS 3800. .5140
/gene="infB"
/note="might be a partial remnant"
/codon_start=1
/transl_table=11
/product="initiation factor 2"
/protein_id="BAC76102.1"
/db_xref="gi:30409153"
/translation="WGCWGVYRKCMANFIRKRSSTIRMDKCDKEKILISVGTIAD
LAKQKKVTEFIRAFQGVIRILNIVBELNQLQGGKGMWQINYEDETDVEVA

```
DYARDDEDEMEIAVASAVERLEQAPIMSNHWSIAIKPGVFLQMPKILKALANTLOIT
DKDVNHPKRAPIIVVESKDKKXILGIIITLILRKSKLXGVDITBCKIRAMQINQREI
KEVKAIGEIVHSGFNQVSLGNTLITSAKKGEKASNFQTNLQKCHPKNLVRSB
SSALAESVQYFLEENKQOLSLAEVSIIMQDVKLQKTSILIVAMQTCVMNTWNPV
PVIKIEANQFEDLLAOLBSRYKVDRIIVSKALVQVFGSIIAGCVLQGEFETGGE
YEIMRNQYQAAARLEIKHKOAVKEVAPQECGIFKKDWTCMVQVDEIVLNVN"
complement (5140. .5212)
/gene="crnt"
/complement (5140. .5212)
/gene="crnt"
/product="tRNA-Thr"
/feature="codon recognized: ACA"
/feature="anticodon (pos:5177. .5179, aa:Thr)"
5213. .5785
/gene="trps4"
5213. .5785
/gene="trps4"
/codon_start=1
/transl_table=11
/product="30S ribosomal protein S4"
/protein_id="BAC76103.1"
/db_xref="GI:30409154"
/translation="KSRYLGRVRIIRRLGLIPATNKSPIKRTGVEGSEAHKTRKS
EYAVLOEKQKQYIYGITNNQARIFRQAKSRASSTGIELKMLERLDRHVRAG
APTLPAARQLVNHGHVKNQVITLASFQCVNHIIEVKAPSSAQLPYLOVENQF
VMIOPEKDMIAFRVNEILVVEYTRVGA"
complement (5788. .6561)
/gene="ycf80"
complement (5788. .6561)
/gene="ycf80"
/codon_start=1
/transl_table=11
/product="ORF257"
/protein_id="BAC76104.1"
/db_xref="GI:30409155"
/translation="MHDIYIEVRLTEDEPKRMOLLQMFVSPVLEHLLSRIGOSV
CIRSPESRIDQESYSPKQOLYHONLBSLAEACIYVYVNPKEILLISROHE
LBYKAPFLMYFDHIFRNPISYLAFHFDPAHNTKSVLNPTEKDKSFEVQSG
LFSFLKNTDLKVLVPLRLQCKMSFEGTIPVFRVQWQHEWLFPSKDAQQFPI
QLSCKITVDVLSLVSLDLDRQARCTLVPLSLNOSL"
complement (6539. .6766)
/gene="rpl28"
complement (6539. .6766)
/gene="rpl28"
/codon_start=1
/transl_table=11
/product="50S ribosomal protein L28"
/protein_id="BAC76105.1"
/db_xref="GI:30409156"
/translation="MSRVCLPLIKKQNAAPVYTSHKRNHRQGVNIOKKRIFINNE
VVLKISTAKIKLKSXTKSTFSNARLILH"
complement (6763. .7071)
/gene="crxm"
complement (6763. .7071)
/gene="crxm"
/codon_start=1
/transl_table=11
/product="thioredoxin type m"
/protein_id="BAC76106.1"
/db_xref="GI:30409157"
/translation="MLHIDELFENEVLQSEKLVVDFWAPCGPCRMIGILEINAK
BNLKAVQVNDENENLATFYGISRIPILMFKKQGVDTVIGAVPSKSLIHTNKL
"
7504. .8970
/gene="rbcl"
7504. .8970
/gene="rbcl"
/codon_start=1
/transl_table=11
/product="ribulose biphosphate carboxylase large chain"
/protein_id="BAC76107.1"
/db_xref="GI:30409158"
/translation="MAQSVQERTRLKKNRYSGVIPYAKGVMVNPVYKQDMLALF
```

```
RVDPQGVDPLEIAAANVAGESSATATVWUTDLTACDVYRAKAYVDPPTAPDOYF
VIAIYIDILPEBSILANLTAISITGNVFGKATKARLEDMRIPESYLTQFQAPGVI
VERERLNKGRPLUGCTVSPKIGLSGKNIGVYBOLKGLDFLQDENTINQPRRM
RORLYLVQVRAAASGEVGSYLVNTVPAAEBCYKRAEFAKEVGSVITMIDLVIG
YVAIQTMAWSENNMILHILHRAGNSYVSROPHQINFRVISKRMRMAGVDIHAGTV
Query Match 34.5%; Score 195; DB 8; Length 14987;
Best Local Similarity 65.0%; Pred. No. 7.6e-44;
Matches 319; Conservative 0; Mismatches 170; Indels 2; Gaps 2;
QY 48 CATGAACACACCCCTCTGTTTGTAGTTGAAGATGAAGCCGAGTCTAACCCGATTGC 107
DB 11049 CATGAACACTACTTTATCTGTTATGATGAAGATGAAGCCGAGTCTAACCTGCAATTGC 10990
QY 108 CGGACTATTTGCGCGCGCGGCTTTTACATGAGAGCTGGCGGTGGCGGAGACA 167
DB 10989 GGGCTTATTTGCGCGCGAGCTTATATGAAAGTTTACAGTACGACCTGCGAGACA 10930
QY 168 GGGGAGCGTTTCCCGCATCACATGCTGCTGCGGGAGTGAACACCATCGAACT 227
DB 10929 AGCAGGTATATCTAGAAATTAACATGCTGCTGCGAGAGATGAACCATGAGCAACT 10870
QY 228 GACCAAGCACTCTACAGTTGTTAAGTATTAAGTACAGACATCACCGAACTCC 287
DB 10869 AATGAAGCAATATATTAACCTGATACCATTTTACAGTGAAGAAATTTGACACAAATACC 10810
QY 288 CTGTGTGAAGAGGAATGATGCTGTGAGGTGAGCGCAATGCGCTTACCGAGCGGA 347
DB 10809 TTGTGTGAAGAGGAATATTTTATGAGGTGAGCGCAATGCGCTTACCGAGCGAGGA 10750
QY 348 AGTGATTGAGCTAGCCAGATATTCGGGCGCGCATTTGGATATTCGGAAGACACCT 407
DB 10749 AATTTAGATATGCGCAATATTTTTCGCGAGGTGGATATATGACGAGCATAT 10690
QY 408 CACCATGAA-TGATGGGAGACCGCGGTAAATGTGAGCAATCTCCAGATGTTGAC-CA 465
DB 10689 CAGTTAGAAAGTACCGGTGATCTCGAAGAAATGTGAGACTCGAACAATTTTAAAGCA 10630
QY 466 AGTGCATTAAGAGGTGCTCGAAGCGGCAAAATGCTTTGGTGGCGGAATCCGGCGT 525
DB 10629 ATTGGTATATGAGAGTACCGCGGTACAGAAAAATTTGATTAACGATCATGAAAGT 10570
QY 526 CAAATCGGAAT 536
DB 10569 TAAACGAGAT 10559
RESULT 9
PUB38804/c 191028 bp DNA circular PLN 27-MAR-1998
LOCUS PUB38804
DEFINITION Porphyra purpurea chloroplast, complete genome.
ACCESSION U38804
VERSION U38804.1 GI:1276652
KEYWORDS chloroplast Porphyra purpurea
SOURCE Porphyra purpurea
ORGANISM Porphyra purpurea
Bacteria; Rhodophyta; Bangiophyceae; Bangiales; Bangiaceae;
Porphyra.
REFERENCE 1 (bases 1 to 191028)
AUTHORS Keith,M.E. and Munnelland,U.
TITLE Complete nucleotide sequence of the Porphyra purpurea chloroplast
genome
JOURNAL Plant Mol. Biol. Rep. 13 (4), 333-335 (1995)
REPERSE 2 (bases 1 to 191028)
AUTHORS Keith,M.E.
TITLE Direct Submission
JOURNAL Submitted (17-OCT-1995) Michael E. Keith, Marine Biology Section,
NRC Institute for Marine Biosciences, 1411 Oxford Street, Halifax,
Nova Scotia B3H 321, Canada
FEATURES
1. 191028
source
/organism="Porphyra purpurea"
/organelle="plastid:chloroplast"
```

```

gene
CDS
/mol_type="genomic DNA"
/strain="Avonport"
/db_xref="taxon:2787"
181..1053
/gene="chlL"
181..1053
/gene="chlL"
/codon_start=1
/product="protochlorophyllide reductase iron-sulfur
ATP-binding protein"
/protein_id="AAC08073.1"
/db_xref="GI:1276653"
/translation="MKLATYGGKGIKSTCSNISVALSKRGKYLQIGCPKIDSTF
TUNGLFPTIDTLOAKDYHEDVWEVDYTGAGVDCVENGGPPAGAGCGYVGE
TVKLKELNAPDEYDILFDVLDGVVCGFAPLNVDYCLITDNGDALFAANRIA
ASVREKARTSLRSLAGLVGNRTDKDLIDKYIDCVMPYLEVLPLEIDIRVSRVSKT
LFEMASDNLAVVCDYNYNIADQLIARPEGVVPKADRELFSLSDFYLNPKNTT
KSTADEEELDMV"
1095..2402
/gene="chlN"
1095..2402
/gene="chlN"
/codon_start=1
/product="protochlorophyllide reductase chlN chain"
/protein_id="AAC08074.1"
/db_xref="GI:1276654"
/translation="WSIVSDALTFECETGNYHTPCPSVSWLYQKIEDSEFLVIGT
KTGCVFLQAMGVMIFAEPRYAMAELEBGDISAKNDYGLRLQLQIKDNPVIF
WIGCTYEIRKMDLBSIAFLKLEIRVIVARANGLDYAFQGEDTVLAAMAQCPG
HHYIENTNHVPLVLEGLDPDVLTQVLEIKQGIIVSGMLPSKRYTELPIQGGY
VAGNPFISRTATTLRRRRKTKIGAPFIPDGPRAMTEKISVFSIKPEGLAREK
DIWDSLEDYALIRGKSVFPMGNLLEVLARFLIRCGWTYVEIGIPMDKRYOAEI
ALLKSTCDMNMVPTIVEKPDNYNLDRIPLKPDVITGMANPLFARINIKMS
VEFTPAQIHGFNARDILSLVTRPLRNLISLELDWDTYSKQ"
complement (2454..3929)
/gene="ORF491"
complement (2454..3929)
/gene="ORF491"
/codon_start=1
/protein_id="AAC08075.1"
/db_xref="GI:1276655"
/translation="NNFIQDLRLLRKSRPIILNTRERELEYIIKQFLCLNEXY
YCMDFVGYTSNPNNGVAKNPNFLALFEIKLDPYALFLIKKDFSPFISLIGIRK
LRMLSOIKTOSKNIIIVGCKNVIPLYLSDVITVDLPNLSEIKETFRSLGIRG
VLIDLELNNAKSCGSLVERIRKTIITITIAOYGOILSRSLPIIEKROLINORLU
ERYPKYAKNEDIGDLALSKMLKRSRSTKOSFNYGIPTRKGLLNGIQGSKRITA
KATANDWLPLELDIGKLFGLVGSBSRMRKMTIAGLSPCVIMIDIKASNL
YSGDSGTSAFVGGTPTWLSERTVFPVAVANTQNLPEMLKRGRTDLPFLDLP
NMERBLFIQIHARIPKSWQNYNIKQSLCNKFSGAIEQAIYESMHTAFSEERE
FSTEDIKIALECFIPLAYTDKEQVESLQAMAGNRAVNASL"
complement (4047..4187)
/gene="trpJ34"
complement (4047..4187)
/gene="trpJ34"
/codon_start=1
/product="50S ribosomal protein L34"
/protein_id="AAC08076.1"
/db_xref="GI:1276656"
/translation="MTKGTIQGSKRKIIISGRPAKMKPSGSLINERRRGRKKIM
AS"
4333..4854
/gene="ycf37"
4333..4854
/gene="ycf37"
/feature="hypothetical chloroplast ORF 37."
/codon_start=1
/protein_id="AAC08077.1"
/db_xref="GI:1276657"
/translation="MLVTLPIFYLSITLPLTFSWLISQOLKTIPLBSQPYFLDK
SKNDELIGEDGFAFSKVCIAKRYFTRAIIESHLVAKNSLSESPENAAIADLYNMG
FIYKADQCKLAKNFYERAIENVGNVVALNNAIKIYEDTKNIKAKAALYDKVLIAK
SNKIANTHKKEIK"
complement (4889..4962)

tRNA
CDS
/gene="trnP"
complement (4889..4962)
/gene="trnP"
/product="tRNA-Pro"
complement (5020..5544)
/gene="ORF174"
complement (5020..5544)
/gene="ORF174"
/codon_start=1
/protein_id="AAC08078.1"
/db_xref="GI:1276658"
/translation="MTFMKTFPQNSTVGSYENKPKLFLDKTPDKIELKQVYLSNT
KNINLYLEOLCDSGVKRPILKKVIALKHSIIISLQKNSSTRLVGAFARATSN
GFNRTIDVAVIHPPQGLGKVVVHQLIKQLQOAEISTITTLFAEPVVISFYKLGSI
KDPDGVGMFWPP"
5807..6364
/gene="psaF"
5807..6364
/gene="psaF"
/codon_start=1
/translation="psaF"
/transl_except="pos:5807..5809,aa:Met)
/product="Photosystem I reaction centre subunit III"
/protein_id="AAC08079.1"
/db_xref="GI:1276659"
/translation="MKKATFTICLAVLAVLSNPVNAEVAAGLIPCKDAAFNKRVN
SVKTLQARIAKVANDNPALANKOIKETKTRFATYGRAGLGGDGLPHLISGRMS
RAGDFVFPGLFLYITTCMGVWGRGILLVAKSKPTBKETIILDVLAIVFMSSGFAM
PLAAMQEPSSGQLIAPNDITVSR"
6377..6505
/gene="psaJ"
6377..6505
/gene="psaJ"
/codon_start=1
/product="Photosystem I reaction centre subunit IX"
/protein_id="AAC08080.1"
/db_xref="GI:1276660"
/translation="MNNNFTKYLSIAFVIGLVMTPTAGFIIELNRFPPVLYYL"
6660..7145
/gene="apcd"
6660..7145
/gene="apcd"
/codon_start=1
/product="Allophycocyanin gamma chain"
/protein_id="AAC08081.1"
/db_xref="GI:1276661"
/translation="MSLVSQIINADDELKPYTLIGLQSIONLYITGSNRPRIATIR
DKKEITIQKSKQIFQALPBYIARGNAAGSRKSLCPRYGVLRLIITYGVLAGDD
SIETIGIIGVREMTNSLGVPIIGMLDIQCLKASLEMLGDDAKIISPFDYIIRGM
S"
7473..8453
/gene="fabH"
7473..8453
/gene="fabH"
/codon_start=1
/product="beta-ketoacyl-acyl carrier protein synthase III"
/protein_id="AAC08082.1"
/db_xref="GI:1276662"
/translation="MGVHILSTGSSVNFVSVENQFEDIETSDHWISTRTGKRRHL
ABSSTLRKLAABANDALSKASINADIDILIAATSPDLFSSASQLQAEIATGS
TAPDITACSGFIIALTAYAGFIOAGSYNKVLVAGDPMRSMWIDMSRSCILFGDGA
GAVLIGBSINSIIGFKLCTDGRNLSHQLMNSPSDQCGELTTVPRGRYDSITMNMK
EYKFAVQVPITVTKNCLDNVNISITDEYDVTILQANRIRLIEATILSLFSKTIIN
LENYGNTSASISIPALDEAIKERKIQGQVVVLAGFAGLWGAIVLMQ"
complement (8465..8536)
/gene="rRNA-Leu"
complement (8465..8536)
/gene="rRNA-Leu"
8735..8854
/gene="psbX"
8735..8854
/gene="psbX"
/codon_start=1
/product="Photosystem II protein X (4.1 kd)"
/protein_id="AAC08083.1"
/db_xref="GI:1276663"

```

gene

RESULT 10

gene	1332.2813
	/gene="rbcl"
CDS	1332.2813

```

1. 6982
/organism="Galdieria sulphuraria"
/organella="plasticid:chloroplast"
/mol_type="genomic DNA"
/strain="UREX 2393"
/db_xref="taxon:130081"
complement (<1. .407)
  gene="matK"
complement (<1. .407)
  gene="matK"
  complement (<1. .407)
    gene="matK"
    /codon_start=1
    /product="maturase"
    /protein_id="AAP81678.1"
    /db_xref="GI:8925946"
    /translation="MFKHGSIFPLDKKKFLFKYLLMTHTPSIGARVPTENMYSKRIA
RADLLNYIVDKKSVDFYALKKFLINKLINLNNIAKRYAIDLFTYDKLTIAKK
REAYNNFRYKRYFLADRLKSHITFYEGFPL"
    complement (522. .860)
    /note="ORF339, hypothetical 12.5 kDa protein"
    /codon_start=1
    /product="unknown"
    /protein_id="AAP81679.1"
    /db_xref="GI:8925947"
    /translation="MMNTYNITKSLADHLNLSIKLNSDEKLYIRIGIKGGGSGLSY
FMTEKKSNSIDNDLVYVDNYNPTLVCDKNSILYLGISLDYSSSLIDGFRFLTPNAK
OTCGGKSGFS"
    complement (877. .1095)
    /note="similar to L28"
    /codon_start=1
    /product="50S ribosomal protein"
    /protein_id="AAP81680.1"
    /db_xref="GI:8925948"
    /translation="WPDIMRKQQLFKTSNTNANVSHSNKSRLQFNLQGGKWS
SLKKWIKLIRSTWMLKLNKKNIIDVFL"
    1332. 2813
    /gene="rbcL"
    /gene="rbcL"
    1332. 2813
    /gene="rbcL"
    /product="ribulose-1,5-bisphosphate carboxylase/oxygenase
large subunit"
    /protein_id="AAP81681.1"
    /db_xref="GI:8925949"
    /translation="MSQSSSEESKVSQETRIKNSRYSGVLPYAKGYMNDPYQKQTD
VLAIFRTLPQGVDPLEAALAAAGVSGSTATWTVVMTDILTADILYRAKAKVQVQPN
PEQVIFREILDLPEGSGTANLTASITGVSGFAYVALTEEMRLPFAVYIKTCQP
AGVILREAILDKRGPILGCTTKPKLIGSGNVRVYVALKGLDFVDDNNISO
PPMKRRRYLPFWAYNKKAAATGVEKGYHLNVTATATBEVNTAAQLAKLGSYITMI
DVLGVYATLADGADNDIMLHRAQNSSTYSKQKHGNNFRVTCMKMRAGVVDHI
DAGVGVATIRRGYKDTLLPKLERNLQDELFPDMDWALRCPMDVPSAGGTHA

```

gene 2859. 3278
/gene="1bcs"
2862. 3278
/gene="1bcs"
/codon_start=1
/transl_except=(pos:2862..2864,aa:Met)
/product="ribulose-1,5-bisphosphate carboxylase/oxygenase
small subunit"
/protein_id="AA081682.1"
/db_xref="GI:8925950"
/translation="MRTIQGFSSFLPDLTDEQIKKQIDYMISKLAIGIEXTNDIHER
NSFMWGGPLFLEVDPAPVLEINACRKAQSNFYIKVGFSSSRGGIESTISPIVNR
PHEGPFNIRROEDRSRISIKYSIQAYETKPEDQRY"
3391. 4308
/note="similar to CfxQ/CfxX protein"
/codon_start=1
/product="unknown"
/protein_id="AA081683.1"
/db_xref="GI:8925951"
/translation="MNVLTETFEYKENINENTLVNIOEYKTKIQEVLNOLDDELIG
LTPVNRKIEKTAALLIDKLRKMLPGANGHMSFTSGDGTGKTVAITMAAILVGL
GCIYIKGHLITATRPDLDVGYQIGHAPKPTKVLKRAMGVFLDEAVLYKPDNERYX
GSEALIEILQWENQRNDVVLVLAGIKQKMRPFIQSNPGLSRRITTNVDPDIYAEFL
LEIGIMLEEQOQYLTBEGKVLLEYIKRRTQHPANARSIRNAIDRRARQNRRLR"
complement(4506..5051)
/gene="ahas"
complement(4506..5051)
/gene="ahas"
/gene="ahas"
/codon_start=1
/product="acetylhydroxy-acid synthase small subunit"
/protein_id="AA081684.1"
/db_xref="GI:8925952"
/translation="MKHTLSYLVDESGVLRISGLFARGFNIDSIAVGPAEYAGIS
RTVMYRGNRIRIOTIKHLYLVNVLKIDITNIPCVERTKLIKYTGNNKNEILL
ETANFRANVVDLSFTILLETGDPGKIALEKULEKYGICETARIGKIALITRESKI
NNEILRLNINNDKIRGGGT"
5201. 5746
/gene="if3"
5201. 5746
/gene="if3"
/codon_start=1
/transl_except=(pos:5201..5203,aa:Met)
/product="translation initiation factor IF-3"
/protein_id="AA081685.1"
/db_xref="GI:8925953"
/translation="MLEKSKQNSPDKNNKFIINNKISPIVRIIDENNYOLGIKIEIE
AKIAEKKLDVLINDKREPVVRLLIDYGRKTFQOEKSEPAKKHQVYKIKRKR
YKIEHDHYQVRINQAIKPKLGNKVKISITFGRELTQIDLAKEULEKIHSLSLAE
HEVFPDSEKSNYIMLPKKT"
5786. 6505
/note="ORF720; putative ABC transporter; ATP-binding
protein"
/codon_start=1
/product="unknown"
/protein_id="AA081686.1"
/db_xref="GI:8925954"
/translation="MKILLEKICFYRRKCLIDNLHMTFSQYEIILGIPNGTGKTSIS
FKLITISAPDKGKVLNGLDITVPMHQRALGICYPQESVFRULVENCILKALMA
EMNGYDAEISSIKSFYSFOIDKTLIFPGSYISGGRRKLEIARAVIIPKPELLML
EPAGIDPISISIELOIIVNLNNGIGILITDHNVRETLKITHRSYIILSGNIAASG
TROLINKARVESYLGNNFTL"
6553. 6691
/note="ORF429"
/codon_start=1
/product="unknown"
/protein_id="AA081687.1"
/db_xref="GI:8925955"
/translation="MFTITDLYSLIOLILGPFVSSAPFTSLSGVLPFLVRRVVE
LPKINGLIKLAKSPFTIFLSPPLASLPLATITPNQNALNKKELIGFRSLGINSYKMLE
STIIFGTLAISISFFINDLIVPSNRRANELLRESKENTL"

ORIGIN

	Query March Best Local Similarity Matches 307;	31.3%; 62.8%; 0;	Score 177; Pred. No. 7.5e-39; Mismatches 180;	DB 8; Length 6982; Indels 2; Gaps 2;
QY	49	ATGAAACACACCCCTCTCTGTTTAAAGTGAAGTGAAGCCGAGTGTCTAACCCCGCATTTGCC	108	
Db	5051	ATGAAACATCATATATCTGTTTAAAGTGAAGATGAATCCGGTGTTTAACTCCGATTTCT	4992	
QY	109	GGACTATTTCGCCCGCTGGTTTAAACATTTGAAGCTTTGGCGGGGTGGCGGAAACAG	168	
Db	4391	GGATTTATTTCTACACGTGATTCATATAGATAGCTTACAGATGTGCTCTGTGAATAA	4932	
QY	169	GGGAGCGTTTCCCGCATCAACATGGTGGTCCCGGGGATGAGAACACATCGAACAACTG	228	
Db	4331	GCAGTATATCAAGAAATTACATAGTATACGGGGTATATAGAAATTATAGAACAAATT	4872	
QY	229	ACCAGCAACTCTACAAAGTTGGTTAACGTAAATTAAAGTAACAGACATCAACCGAACTCC	288	
Db	4871	ACAAACATCTTTTATTAATTAGTAAATGTATTTAAATTCAGATATACAAATATTTCCA	4812	
QY	289	TGTGTGAAAAGGAATTGATGCTGTGTGAAGTGAAGCCCAATGCCCTTACCCGACCGAA	348	
Db	4811	TGTTAGAAAGAGATTTAAATGCTATATAAAGTGAATACTGGAGAGAAATTAAGAAATGA	4752	
QY	349	GTGATTGAGCTAGCCCAAGTATTCGGGCCCCGATGTGGATATCTCCGAAGACACGCTC	408	
Db	4751	ATATTGGAAATGGCAATATTTTTAGAGCACGTGTGTATTTATCAGAACTTTTATT	4692	
QY	409	ACCATCGA-ATGTTG3GGGGAACCCGGGTAAATGGTAGCAATCTCCAGATGTTGGCCAA	467	
Db	4691	ATTTTAGAGTAACAGAGAGATCTGGCAAAATGCTGTATATGAATAATTATAGAAAA	4632	
QY	468	T-TGGCATTTAAAGGCTGCTCGAAGCGGCAAAATTGCTTTGGTGGGGAATCCGGGCTC	526	
Db	4631	TATGTAATTTGTGAATTTGCTAGAACAGGCAAGATTGCTTTACAGAGAAATCAAAAATA	4572	
QY	527	AATACGGAA 535		
Db	4571	AATACTGAA 4563		

	RESULT 11		
LOCUS	BX572098	346683 bp	DNA
DEFINITION	Prochlorococcus marinus K17913 complete genome; segment 47.		circular BCT 23-SEP-2003
ACCESSION	BX572098 BX548175		
VERSION	BX572098.1 GI:33640689		
KEYWORDS	complete genome.		
SOURCE	Prochlorococcus marinus str. MIT 9313		
ORGANISM	Prochlorococcus marinus str. MIT 9313 Bacteria; Cyanobacteriota; Prochlorophytes; Prochlorococcaceae; Prochlorococcus.		
REFERENCE	1 (bases 1 to 346683)		
AUTHORS	Rocap,G., Latimer,F.W., Lamerdin,J., Malfatti,S., Chain,P., Johnson,N.A., Arellano,A., Coleman,M., Hauser,L., Hess,W.R., Albright,D.I., Land,M., Lindell,D., Post,A.E., Regala,W., Shah,M., Shaw,S.L., Steglich,C., Sullivan,M.B., Ting,C.S., Tolonen,A., Webb,E.A., Zinser,E.R. and Chisholm,S.W. Genome divergence in two Prochlorococcus ecotypes reflects oceanic niche differentiation Nature 424 (6952), 1042-1047 (2003)		
JOURNAL	Nature 424 (6952), 1042-1047 (2003)		
MEDLINE	22825698		
PUBMED	12917642		
AUTHORS	2 (bases 1 to 346683) Latimer,F. and Rocap,G. Prochlorococcus genome consortium Direct Submission Submitted (03-JUN-2003) Submitted on behalf of the Prochlorococcus genome consortium, the DOE Joint Genome Institute, Production Genomics Facility, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA, and the Genome Analysts Group, Oak Ridge National Laboratory,		
TITLE			

1060 Commerce Park Drive, Oak Ridge, TN 37831, USA;
larimerf@ornl.gov
location/Qualifiers
1..346683
/organism="Prochlorococcus marinus str. MIT 9313"
/mol_type="genomic DNA"
/strain="MIT9313"
/db_xref="taxon:74547"
complement(347..637)
/locus_tag="PMT0970"
complement(347..637)
/locus_tag="PMT0970"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="possible Herpesvirus VP23 like capsid protein"
/protein_id="CAE21145.1"
/db_xref="GI:33640690"
/translation="MKIKHSIRHIFESCTPEPLICIRKYSKOTLISCDLVNDIEPLA
WIAGFVVDLIFSNKRPDPLILGATFIVLALSHVSPSDVCKASIMQHALS"
620..1015
/locus_tag="PMT0971"
620..1015
/locus_tag="PMT0971"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="possible bZIP transcription factor"
/protein_id="CAE21146.1"
/db_xref="GI:33640691"
/translation="WDFDSCKPSTSRSGDYYRPAIVITTVRSEILMASSCGGLE
INDKSRPEVYNGPFGYITPRDVRVYIPDSKYKAQERQNAQVRLARRAEILSY
VDQEKQIMEDQPSQDEVPDKGLAAKT"
1478..1699
/locus_tag="PMT0972"
1478..1699
/locus_tag="PMT0972"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="possible DNA polymerase (viral) N-terminal dom"
/protein_id="CAE21147.1"
/db_xref="GI:33640692"
/translation="YPMOQLRRLLEFLKPRVLAQAEFFTPMOSKAGIDICILRW
LSDVPEALADDQLKQVLDQARWLYN"
2436..2690
/locus_tag="PMT0973"
2436..2690
/locus_tag="PMT0973"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="possible vinculin family"
/protein_id="CAE21148.1"
/db_xref="GI:33640693"
/translation="MQGMYGDERWLRHMAALFSCFRWLGWMTGLFSKRDLQVEFKP
PKEVTFKPDDELSPISRDDEFKRLLEQTLAESPSQED"
complement(2865..3230)
/locus_tag="PMT0974"
complement(2865..3230)
/locus_tag="PMT0974"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="CAE21149.1"
/db_xref="GI:33640694"
/translation="WQMDRNVLGEPLIEICDPTGMYRGCKCRTHDGLGQHTVCCV
ISEQFLYSKAQNDLSTPIQYFGSGIKPDHWCVCASRWKQAYDDGMAPNIRLEAT
NIEVLAIDLILKQYSHT"
complement(3351..4007)
/locus_tag="PMT0975"
complement(3351..4007)
gene
CDS

sig_peptide
/locus_tag="PMT0975"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="possible Ribosomal protein L36"
/protein_id="CAE21150.1"
/db_xref="GI:33640695"
/translation="MRKPSIDPIFSAAGILLCTSTISAKKDVIGTYSYKCR
EGIMKSKTIGRVNDKMTITERLKLSTGRVYVNDIRKSCQSSLISKAGATGA
DCELASDFPSNWMQIDYASRFNAHGNKHTIFAIRYDGSRLCLSKNNFRDLPLET
PGRSLFLHNIERQADSIILEYFHPNGWGYEKVRYLPLSPERPSPFVDSWIHOR
"
complement(3918..4007)
/locus_tag="PMT0975"
/note="Signal predicted by SignalP 2.0 HMM (Signal peptide
probability 1.000) with cleavage site probability 0.694 at
residue 30"
4938..5267
/locus_tag="PMT0976"
4938..5267
/locus_tag="PMT0976"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="possible Protein of unknown function DUF42"
/protein_id="CAE21151.1"
/db_xref="GI:33640696"
/translation="MGFGALQVGRINQAGIDAPLPIAVGNSVFFIDALYANPSF
YDNVSSILNTVPMKACPFGRKSGTLMLPFYKVGSGMGLLEDQGSRAVYQLISAG
PNAVND"
complement(5257..7404)
/locus_tag="PMT0977"
complement(5257..7404)
/locus_tag="PMT0977"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="ABC transporter, multidrug efflux family"
/protein_id="CAE21152.1"
/db_xref="GI:33640697"
/translation="MTTNPGRPALQARALEELLFSYTPENEGEIEELRLGFCVAKL
GRQKLOPLSGASIKELNFDNHRKIEALPDNQAEYPLVVFDFEHNPFALYR
HGGQWMLPDDQESHRLQDSNKLKATAEVYPSIPVYGLAIVRAPAFELGALT
ALVITSAVYMTLSTPMLTNTVSRILPQNDQNLIFQGLTVVITIVIGSAVEPLQS
LMLRLSTADLRLOSAVMDRLRLPMSITSYTTGDLASRSALISQRLIGNVLIS
TLISLFAISFEFLMFEVYDNLAWASFTLVAAICLIMITVRSIQRLPLESGAEI
TNPSLQAVMGQPIRSAGSEPYLVRWLREIRVYASLQRSVYSDAMEQYGLTVPL
ASLPMFAVVAOVVNSPSEINLTQVAFISNAAFGNGALTGAINLIANVAGRAA
VLMRREPVLVADVEKGYQPEALVHSHSGEPLRIGIYSEFPSSROLNLSLTIPAG
OHTAITGASGCGKTVLVMTLGFISPGCEILVDSIPITQALAIRYRQLGVMTGAR
LNTGSIYEYCQGVSRSEVYQALQOALADERVAMPQVLETLISDSGNSGQVQ
RIARALITQPRVLIIMDEATSLDNRQERTQTQIKELGMTRISIAHRLSTIQADQ
IVLERNQKPEVMSGKMDLKSGLIASLASH"
order(complement(6868..6933), complement(6757..6822),
complement(6523..6588), complement(6451..6507),
complement(6187..6252), complement(6076..6111))
/locus_tag="PMT0977"
/note="6 probable transmembrane helices predicted by
TMHMM2.0"
complement(7401..9617)
/locus_tag="PMT0978"
complement(7401..9617)
/locus_tag="PMT0978"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="ABC transporter, multidrug efflux family"
/protein_id="CAE21153.1"
/db_xref="GI:33640698"
/translation="WKLQKRVKQAQRLQYAEACGASALATILAYRGLVEISELR
IACGVNRGSNAKQVLLKAPRYGLAKYRQSGEPLSEGRPCVFGNRPFLVYRG
RNDRAVPLSDPAQKTRVEMBEFLDKFTGIYVLEBPGRPGRGSEKRLMSPKXKL
PGRSLQDLTLCVASHQVFGIIVAGFTSGFIDNQLQNERLYFGIPIITMLPLMSVGLWL

VERSION AF022186.2 GI:6466296
 KEYWORDS
 SOURCE
 ORGANISM
 chloroplast *Cyanidium caldarium*
Cyanidium caldarium
 Eukaryota; Rhodophyta; Bangiophyceae; Porphyridiales;
 Porphyridiaceae; Cyanidium.
 REFERENCE 1 (bases 130696 to 132364)
 AUTHORS Vogel, H., Fischer, S. and Valentin, K.
 TITLE A model for the evolution of the plastid sec apparatus inferred
 from secX gene phylogeny
 JOURNAL Plant Mol. Biol. 32 (4), 685-692 (1996)
 MEDLINE 97134960
 REFERENCE 2 (bases 1 to 164921)
 AUTHORS Gloeckner, G., Rosenthal, A. and Valentin, K.
 TITLE The structure and gene repertoire of an ancient red algal plastid
 genome
 JOURNAL J. Mol. Evol. 51 (4), 382-390 (2000)
 MEDLINE 20496959
 REFERENCE 3 (bases 46857 to 47851)
 AUTHORS Valentin, K.
 TITLE Direct Submission
 JOURNAL Submitted (22-MAR-1996) Institute for Plant Physiology, Justus
 Liebig University, Heinrich Buff Ring 58-62, Giessen 35392, Germany
 4 (bases 28701 to 75580)
 AUTHORS Gloeckner, G., Rosenthal, A. and Valentin, K.
 TITLE Direct Submission
 JOURNAL Submitted (02-SEP-1997) Department of Genome Analysis, IMB Jena,
 Beutenbergstr.11, Jena 07745, Germany

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT	FEATURES
Direct Submission	Gloeckner, G., Rosenthal, A. and Valentin, K.	Submitted (18-NOV-1999)	Genome Analysis, Institute for Molecular Biotechnology, Beutenbergstrasse 11, Jena 07745, Germany		source
Direct Submission	Vogel, H., Fischer, S. and Valentin, K.	Submitted (18-NOV-1999)	Institute for Plant Physiology, Justus Liebig University, Heinrich Buff Ring 58-62, Giessen 35392, Germany	On or before Nov 22, 1999 this sequence version replaced gi:529651, gi:1240002, gi:2465730.	location/Qualifiers
					1. .164921
					/organism="Cyanidium caldarium"
					/organelle="plastid:chloroplast"
					/mol_type="genomic DNA"
					/strain="RK1"
					/db_xref="taxon:2771"
					join(164808..164921,1..78,462..644)
					/note="direct repeat;"
					repeat unit separated by stem loop"
					339..477
					/note="stem-loop separates direct repeat unit"
					580..780
					/note="similar to yeast mitochondrial origin of replication region in Genbank Accession Number L36902"
gene					986..2041
					/gene="psbd"
mrna					986..2041
					/gene="psbd"
					/product="unknown"
					986..2041
					/gene="psbd"
					/note="Photosystem II D2 protein"
					/codon_start=1
					/transl_table=11
					/product="unknown"
					/protein_id="AAP33026.1"
					/db_xref="GI:6466445"
					/translat="MTTALASGEORGMFDLLDMVKRDRFVFGMSGTLIPCPAYIALGAWTGTTPVSWYTHGLASXSGECGFLTRAVSSPNSMGSLFLFMGPAODPFRWCOQTGIMTFLAHGSGFLIGFCURQCEIRALVCLRPYNALAFSGPIAVFVSVLIVF
CDS					

Accession	Sequence	Position
Db	106592 TCGTGTGAGCGTAACTCATCTACTAGTAAATAATTAAATGCTAGTACTAGTTCCAGACCGAAA	106533
Qy	349 GTGATTGAGCTAACCACGATATTCGGGSCCGCATTTGGATATCTCCGAAGACCGTC	408
Db	106532 ATTTTAAATATTTTGGAACTATTAGAGCTAAAGTAGAGATTATACGAAAATCTGTTA	106473
Qy	409 ACCATCGAA-TGGTGGGGCACCCGGGTAAATAGTAGAATCTCCAGATTTTG-GCCAA	466
Db	106472 GTTTTGAATATGCACGGGATCCAGGTAAAGTAGCTGCAATCAACAGCTCTGCAAA	106413
Qy	467 GTTGGCAATTAAAGAGTGCTCGAATCCGGCAAAATTGTTTGGTCGGGAATCCGGGCTC	526
Db	106412 TTTTGGATTAATACAAATTTGCAAGAACTGGCAAAATAGCGTTTAACAGAACTCTAAAATA	106353
Qy	527 AATACGAATATCTGAAA	544
Db	106352 AACACAGAGTTTGGAAA	106335

RESULT 13	LOCUS	DEFINITION
AE017164	302399 bp	DNA linear BCT 14-AUG-2003
AE017164		<i>Prochlorococcus marinus</i> subsp. <i>marinus</i> str. CCMP1375 section 4 of 6 of the complete genome.

SOURCE ORGANISM	REFERENCE
<i>Prochlorococcus marinus</i> subsp. <i>marinus</i> str. CCMP1375	1 (bases 1 to 362399)
<i>Prochlorococcus marinus</i> subsp. <i>marinus</i> str. CCMP1375	Dufresne, A., Salsanoubat, M., Partensky, F., Artiguenave, F.,
Bacteria; Cyanobacteria; Prochlorophytes; <i>Prochlorococcus</i>	
<i>Prochlorococcus</i>	
1 (bases 1 to 362399)	

TITLE Genome sequence of the cyanobacterium *Prochlorococcus maritimus* strain P-72-22
 AUTHORS Dufresne, A., Salanoubat, M., Partensky, F., Arriens, F.,
 JOURNAL Sci. 200, Natl. Acad. Sci. U.S.A. 100 (17), 10020-10025 (2003)
 REFERENCE 2 (bases 1 to 30339)

REACTURES
JOURNAL
TITLE
DIRECT SUBMISSION
Submitted (28-MAY-2003) UMR 7127 CNRS et Université Pierre et Marie
Curie, Station Biologique de Roscoff, Place Georges Tessier BP74,
Roscoff 29682, France
Location/Qualifiers

```

FEATURES
    source
        location/Qualifiers
            1..30239
                /organism="Escherichia coli"
                /mol_type="genomic DNA"
                /strain="CGP1375; S5120"
                /sub_species="marinus"
                /db_xref="taxon:167539"
            103..1146
                /gene="acsP"
                /gene="acsP"
                /locus_tag="Pro0992"
            103..1146
                CDS

```

```

/locus_tag="P00092;"
/note="PM:1081605; di-iron enzyme CcdI expression is
activated in copper- or oxygen-deficient cells"
/codon_start=1
/transl_table=11
/product="Mg-protoporphyrin IX monomethyl ester aerobic
cyclization protein homolog"
/protein_id="AAQ00037.1"
/db_xref="GI:33237970"

```

/translation="MTTTSQAPASLGRNELPHLDNLTLPRYTTTEPKAAKTDL
 LAKQPEMFMEKEADVILKHPFDKASLERLDELSPEDKAYESYLVRVSVESEGL
 LERKSNFPEKAGQELQGFPTFLARDEARGLGALKTGELINVLPLPKRAAT
 FEPPLSVWVSYSLSEKIGYMYRILINRLKANPQVCAPLDFPEPWCQOENRGGDI
 NMMGCMGKRGFRGLKSLSPFLMTVLTHTLTVCEGSEPEYELIGDPLYDPEVIL
 QNNTKNAFPMVYKFEEDGKFLSRIDILNAFKKMRQNGIKRLALGKTVLLKPF
 TLPMKRTAYVYQG"
 complement (1153. .1515)
 /gene="petf"
 /locus_tag="Pro00993"
 complement (1153. .1515)
 /gene="petf"
 /locus_tag="Pro00993"
 /codon_start=1
 /transl_table=11
 /product="Ferredoxin, PetF"
 /protein_id="AAQ00038.1"
 /db_xref="GI:33237971"
 /translation="VKSXKITVRLRDNRITISKEIPGESILRKEEDGVLPPSCGN
 GCTTCARILISGTTIDOHDIGISKQNGENYGLLCTAKTGRADLETQDBDYELQ
 FGRYIGSKNTGPFDM"
 complement (1538. .1870)
 /locus_tag="Pro00994"
 complement (1538. .1870)
 /locus_tag="Pro00994"
 /codon_start=1
 /transl_table=11
 /product="Predicted protein"
 /protein_id="AAQ00039.1"
 /db_xref="GI:33237972"
 /translation="MTQOQKLYVLLAPIGOLCGNGOLRETISERNRGQGVFWYLS
 ADVLRNIASTEFEAVVAADKRLDWLKLRFGEELSIMDIDELIKQDMDLPPAP
 IKDISIPD"
 1999. .3450
 /gene="clD"
 /locus_tag="Pro00995"
 1999. .3450
 /gene="clD"
 /locus_tag="Pro00995"
 /note="rlD"
 /locus_tag="Pro00995"
 /codon_start=1
 /transl_table=11
 /product="Predicted Zn-dependent protease"
 /protein_id="AAQ00040.1"
 /db_xref="GI:33237973"
 /translation="MKVALTKARSENFLKMTDOIIDLKRNLESLAYGISADAPPE
 IFENIDILARLADPTVTSVNSPIGSGVGVRLKCDGFVSTINSPSEGLKPALDQ
 ALGNLGRNRISIKRFSFGLNVIDYKRAKSGVSLCPSLKESTNKYLETSLIQRG
 PHISIRDSVSKTFOEVIIVASSDGTGRDLRYQSLGNVLALDKQRRSIGRRIGSA
 GKPYLLDMWDEAKETAINBSAEKMLYAKTVASQGMVAVLANFEGVIFHEACGHLE
 TQIQRGTTPEQKLNQOIAKNAVTAVEGLTQGAFGSISMDGEMTOKTVLIENGI
 LKQFLSDPAGALRTGHNRTSGRRQSYAPAAASMRNTYIEKGYPDENLIESIDNII
 YCKMGSGSYAGATQGFNPSVBEGLIKNGKLTDPYKATLILGAKEVMPRLSMCADD
 DLAAFGSGISGANVTAVGQPHIKVDSTVGR"
 3453. .4832
 /gene="pmbA"
 /locus_tag="Pro00996"
 3453. .4832
 /gene="pmbA"
 /locus_tag="Pro00996"
 /note="pmbA"
 /codon_start=1
 /transl_table=11
 /product="Predicted inactivated Zn-dependent protease"
 /protein_id="AAQ00041.1"
 /db_xref="GI:33237974"
 /translation="MTMSNNBSLISEGNLADKLTLSNEIKKMDLGASIGQTSVQ
 IDKGEPIKRSQKSSITIRVWNNKENVGILTSDSAGLRKALNSAFSAFPAKL
 ESNPFSKLQNALPNLQRPITQSLGVKCLAYILEASGKLDKNSIKSLPYNGFAS
 KYEAVYINSDAFHMSSTASLYLVAPAEERKRRSGSISKLAAVAADINIEGIN
 BAAKTRHLNVALEIKTNKYLICFKPEAPFDLTSFNINVASIIDGISLRKNSG
 EOTISAGNSLDEGLHKDNGACTPDEGPTPNTIILDSGILKNIILHSATARTGC
 NPTGHAIGKLVKSDVNTVTKTETTAAYPDLQGMKREEVNIVENHAIHAGICS

SOGSGSLPFDGMLYKNGEQIESATTAAGDIKLLKNIQIEDDNGSTHOGVCHITW
 DELAVTGA
 gene
 /gene="fmc"
 /locus_tag="Pro00997"
 4829..5848
 /locus_tag="Pro00997"
 4829..5848
 /gene="fmc"
 /locus_tag="Pro00997"
 /codon_start=1
 /transl_table=11
 /product="Methionyl-tRNA formyltransferase"
 /protein_id="AAQ00042.1"
 /db_xref="GI:33237975"
 /translation="MNTIFMGCTPFCVPILEKLLKSNHNVLAITVTPDRGRGKNT
 PSPKOKALEENLPITTPVNTKEKDIOAKIKNADIPVYVAFGQILPKSVLKPRY
 GCWNTHASLPRMRGAFTQWSILSGSETEGVGAMABEGLDTAVLLEKLLKULKE
 NAEQSQRLKDLSCNLIIEGILEKXEQSLSKNTLKQEDINTFYARLKLME
 DYLIINMDSGNIKOILGLPNCYATINSKRLKLSIPINNKYDPELFRYDIL
 KSGSINNNGIIDIENIGIIVQTKVPLITVKLEKESQNALIQDTLINSO
 NKLGE"
 gene
 complement(5849..6571)
 /gene="terc"
 /locus_tag="Pro00998"
 complement(5849..6571)
 /gene="terc"
 /locus_tag="Pro00998"
 /note="possibly involved in tellurium resistance"
 /codon_start=1
 /transl_table=11
 /product="Conserved membrane protein Terc"
 /protein_id="AAQ00043.1"
 /db_xref="GI:33237976"
 /translation="MDIASLDYLPFLNGIDQWSEFATLLPFLVLELLISADNAVAL
 ASITRDLSTLEOKSLNIGIATISLPIILLIANTIIKYPVIAQAAVLSIVT
 KPTINPNDESSISNNKSPFKITPELLIADPLASVPIAAVAISDQILVITG
 ALIGITALRPTADLPIRDLDFVQLENAGTAAVAIKLFEYTPNIVIPITPII
 LLGFVFLMGSKRPVPLNSENE"
 gene
 6665..7375
 /locus_tag="Pro00999"
 6665..7375
 /locus_tag="Pro00999"
 /note="conserved in cyanobacteria"
 /codon_start=1
 /transl_table=11
 /product="Uncharacterized protein"
 /protein_id="AAQ00044.1"
 /db_xref="GI:33237977"
 /translation="MOINSELCHVDDEKIIIVITALENOIQAGLQSGSNVHAES
 ALOOLITLKEVOKIOSINSNEAKSLVONSSTKEIEPKSKENSKISKDNYD
 IYDCKVPMDSERISKDIERLGMKREBENPKITTGSRNRKIKSYEDLPII
 ELQMKRGDMKNININYLEKULISENDILRLSNMSTAREYFVDRKFKVNSREL
 GIYELKFSVSLGRELKS"
 gene
 7440..10952
 /gene="mfd"
 /locus_tag="Pro1000"
 7440..10952
 /gene="mfd"
 /locus_tag="Pro1000"
 /note="superfamily II helicase"
 /codon_start=1
 /transl_table=11
 /product="Transcription-repair coupling factor"
 /protein_id="AAQ00045.1"
 /db_xref="GI:33237978"
 /translation="WALKSLITNVESALTEIYRVYKVKKSLINGSGTSAATIS
 SLAKSNRPVIVPTLEASRMSIMQILGMNVLLYPSSEVPSYISIKSEIWA
 OLVSLELANDYRDPMAICTERALQPHLPRPLMENSVISIKGELMDLNSRL
 IKIGYEKTTQEGTMRGDIIDIYAVNNELPIRELFEDNDIXREPDPINQRL
 DEINNRISPVINKITPNLNNELKKNKNGHEIEGTIEELLETOAKDSILMY
 IAKSLICTDKKQCVSHENWNIHRTTEFLMDNTLKNNSINTLNNDKNOLS
 ILDEPKTIEKNDISINQDSPHNSKITTYTNGSKSEILREQKKKYSWLS
 AOPSPASALBSEHCITKPVONNLDISHINLIEFKIPVSKLHGNLDIGVDSQK
 VLLTDKELFGQIVSSTGIFTRKRSSSRSVNKKSGDGVVHRNHGKFLKEK

FIVNNSRDYLLVOYSDGTPLVAADOLSLGRVNSDPSKSPINKLGNWTAKERA
 KKSIRVAIDLKIAHNSNGSEFPDQGMORELEBAFYEATPDQIKAVSEKSD
 MEKSPMDRLVCGDVGFKTEVAIPALFKAITAGQVAIAPTVLAAQOHWRTLDRF
 Query Match 28.6%; Score 161.8; DB 1; Length 302399;
 Best Local Similarity 60.3%; Pred. No. 2.1e-34;
 Matches 302; Conservative 0; Mismatches 197; Indels 2; Gaps 2;
 QY 48 CATGAAACACCCCTCTGTTTACTGTAAGATTAAGCCGAGTCAACCCGATTCG 107
 DB 242767 CATGAGCACTCTTTCAGTCTAGTAGAGATCAATCAGTCTTACTGTAATTC 242826
 QY 108 CGGACTATTCGCCCGCGGTGTTTAACTAGAGACTTGGCGGTGGCGGAGACA 167
 DB 242827 AGGCTATTCCTCCGAGAGATTCACATCAACGCTTGTGTGGCCCTGAGAGC 242886
 QY 168 GGGGACGTTTCCCGCATCACCATGCTGCGCGGGAGTGAACACCATCAACAAC 227
 DB 242887 TGAAGCATTTCTCGCTTAAACATGCTGTTGAAGAGATCATCAACATCCACAAT 242946
 QY 228 GACCAAGCACTCTCAAGTGTGTAAGTAACTTAAGTACAGCATCACCGAAATCC 287
 DB 242947 GACAAAGACCTTGATTAATTAATCAATGATTAAGTATCTTAACAACATTTCC 243006
 QY 288 CTGTGTGAAAGGAATTGATGCTGTGAAGGTGAGCCCAATGCCCTAACCGAGCGA 347
 DB 243007 TGCCGTAGAGAGAAATTAATCTTTAAAGTATCTCAAGTAAAGAAAGGACAAG 243066
 QY 348 AGTGATGCTAGCCAGCTATTCGGGCCCGCATTTGATGATCTCGAAGACACCGT 407
 DB 243067 CATCTTGATTTAGTTCAAGTTTAAAGCAAAAGTTGATGATGAGAGATCTCT 243126
 QY 408 CACCATCGAA-TGTTGGGGGACCCGGTAAATGTAGCAATCTCCAGATGTTG-GCCA 465
 DB 243127 TACATTAGAGAGTTGAGAGATTCACGAGGAACTGTTAGCCCTAGAAAAGCTTTGAAGC 243186
 QY 466 AGTTGCATTAAAGAGGTGGCTCGAACCGGCAAAATTTGTTGTGCGGGAATCCGGGCT 525
 DB 243187 TTATGCATCTCTTGAGTAGAAGCAAGCTGGAAGAGTTGCACTTGAAGGGCCTCAGGTG 243246
 QY 526 CATACGAAATATCTGAATC 546
 DB 243247 TAATACGAGATGTTGAATC 243267
 RESULT 14
 AF041468/c 121524 bp DNA circular PLN 03-MAR-1999
 LOCUS
 DEFINITION
 ACCESSION
 X62348 X14171 X62349 X51511 X14504 X52158 X52912 X56806 X76547
 X62348 X21576 U81044 AF063017
 AF041468.1 GI:3602932
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 chloroplast *Guillardia theta*
Guillardia theta
 Eukaryota; Cryptophyta; Cryptomonadaceae; *Guillardia*.
 REFERENCE
 1 (bases 47701 to 48415)
 Douglas, S.E. and Durnford, D.G.
 The small subunit of ribulose-1,5-bisphosphate carboxylase is
 plastid-encoded in the chloroplast 1-c-containing alga *Cryptomonas*
phi
 JOURNAL
 Plant Mol. Biol. 13 (1), 13-20 (1989)
 MEDLINE
 93357429
 PUBMED
 2562756
 REFERENCE
 2 (bases 18535 to 19351)
 Douglas, S.E. and Durnford, D.G.
 Sequence analysis of the plastid rDNA spacer region of the
 chloroplast 1-c-containing alga *Cryptomonas phi*
 JOURNAL
 DNA Seq. 1 (1), 55-62 (1990)
 MEDLINE
 92119320
 PUBMED
 2132959
 REFERENCE
 3 (bases 43739 to 44938)
 Douglas, S.E. and Durnford, D.G.

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE

Nucleotide sequence of the genes for ribosomal protein S4 and tRNA(Arg) from the chlorophyll c-containing alga *Cryptomonas phi* 90245597

2336372

4 (bases 34539 to 35380)

Reith,M. and Douglas,S.

Localization of beta-phycoerythrin to the thylakoid lumen of *Cryptomonas phi* does not involve a signal peptide

Plant Mol. Biol. 15 (4), 585-592 (1990)

91338697

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE

5 (bases 45872 to 47981)

Douglas,S.E., Durnford,D.G. and Morden,C.W.

Nucleotide sequence of the gene for the large subunit of ribulose-1,5-bisphosphate carboxylase/oxygenase from the chlorophyll c-containing Alga *Cryptomonas F*: evidence supporting the polyphyletic origin of Plastids

J. Phycol. 26, 500-508 (1990)

6 (bases 110917 to 113854)

Douglas,S.E.

Unusual organization of a ribosomal protein operon in the plastid genome of *Cryptomonas phi*: evolutionary considerations

Curr. Genet. 19 (4), 289-294 (1991)

91330343

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE

7 (bases 40675 to 42376)

Douglas,S.E. and Turner,S.

Molecular evidence for the origin of plastids from a cyanobacterium-like ancestor

J. Mol. Evol. 33 (3), 267-273 (1991)

92099311

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE

8 (bases 96123 to 96906)

Wang,S.L. and Liu,X.Q.

The plastid genome of *Cryptomonas phi* encodes an hept0-like protein, a histone-like protein, and an acyl carrier protein

Proc. Natl. Acad. Sci. U.S.A. 88 (23), 10783-10787 (1991)

92073372

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE

9 (bases 106789 to 108216)

Douglas,S.E.

A secy homologue is found in the plastid genome of *Cryptomonas phi*

FEBS Lett. 298 (1), 93-96 (1992)

92183838

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE

10 (bases 42198 to 44153)

Douglas,S.E. and Reith,M.E.

A bohi homolog, encoding a subunit of Mg chelatase, is located on the plastid genomes of red and *Cryptomonas* algae

J. Mar. Biotechnol. 1, 135-141 (1993)

11 (bases 82327 to 84479)

Douglas,S.E. and Murphy,C.A.

Structural, transcriptional and phylogenetic analyses of the atpB gene cluster from the plastid of *Cryptomonas F* (Cryptophyceae)

J. Phycol. 30, 329-340 (1994)

12 (bases 98901 to 114602)

Wang,S.L., Liu,X.Q. and Douglas,S.E.

The large ribosomal protein gene cluster of a cryptomonad plastid: gene organization, sequence and evolutionary implications

Biochem. Mol. Biol. Int. 41 (5), 1035-1044 (1997)

97283757

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE

13 (bases 61067 to 68605)

Leitsch,C.E.W., Kowalik,K.V. and Douglas,S.E.

The alpha gene cluster of a cryptomonad, *Guillardia theta*: A piece in the puzzle of chloroplast genome development

J. Phycol. (1998) In press

14 (bases 1 to 121524)

Douglas,S.E. and Penny,S.L.

The plastid genome of the cryptophyte alga, *Guillardia theta*: complete sequence and conserved synteny groups confirm its common ancestry with red algae

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL

J. Mol. Evol. 48 (2), 236-244 (1999)

99128221

15 (bases 1 to 121524)

Douglas,S.E.

Direct Submission

Submitted (08-JUN-1998) Institute for Marine Biosciences, National Research Council, 1411 Oxford Street, Halifax, Nova Scotia B3H 3Z1, Canada

On or before Sep 15, 1998 this sequence version replaced gi:11396, gi:11297, gi:18103, gi:18281, gi:11383, gi:11407, gi:12539, gi:136730, gi:11300, gi:398949, gi:2661180.

Location/Qualifiers

1..121524

/organism="Guillardia theta"

/organelle="plastid:chloroplast"

/mol_type="genomic DNA"

/db_xref="taxon:55229"

/note="plastid"

/complement(5..86)

/gene="trnY(gua)"

/complement(5..86)

/gene="trnY(gua)"

/product="tRNA-Tyr"

/note="codon recognized: UAC"

/anticodon=(pos:50..52,aa:Tyr)

/complement(181..573)

/gene="rpl19"

/complement(181..573)

/gene="rpl19"

/codon_start=1

/product="ribosomal protein l19"

/protein_id="AAC35594.1"

/db_xref="GI:3602934"

/translation="METILLTYEFVNKSKNLIQLVESHSKTLPLQGLVDTRVIGVSTIEGSKRIQATEGVIIISKNGINLTITIRVQIGIERVYLLHSFKSIETLRKSKIRAKLYIYRTSGKATRLQRFN"

729..3188

/gene="clpC"

729..3188

/gene="clpC"

/codon_start=1

/product="Clp protease ATP binding subunit"

/protein_id="AAC35595.1"

/db_xref="GI:3602934"

/translation="MFERFTEKAIKYIMLAQEARSLGHNVGTEDILIGLEGNGIARVLSMGVNLDAVEKTIIRGSGPVAVIPTTPAKRILELSIEBARLGHTYIEHLLIGIRREGCVAAVLENLADLTAKTVQVIRLGDVAVSATKQTKTPTIEFGSNLQKABEGLDVIGRQKEIEKVIQILGRTRKNPILIGEPVGATAIAEGLAQRLNNRPVPILEDKRVTLIDGLVAGKYRGEFERLKKIIDEIRVANNVILV IDEVHLIGAGAEAGAIIDANILKPALAGEMQICATLIEERKHI EKDSALERFQ PVWGESPVSEETIEILYGRDRKEKHKVI SEBLSAAKPADQYIADRPDKXID LIDEGASRVAKMSOLPPAARELDKELRLIKSDKSAVSOPEETGCLRPMETIKACIAIHSKSGDEENKESVYVTEEDIAIVAMTGI PYNKTRSESKILQNEFTLH GRITGDEAVANVAKIRRAVATKRPKNRPIASFTSGFGVGTETLTALAVYFGS ERMVRLDSSEVIERHTVSKLISPGVYVNEGQLTSSVRRRPYTVLVFDEIENGH PDVFNNLILILEDGRLTDSKRTVPDKNTLLITSNVSGKVEREGGGLGFDISEQT ESQYGRIKALVNEELKQYRPERPLANLDEIIVRQLTKDEVGIAIMLKEVPTRISE KGIQLEVTARFKTHLINEGNPNPYGARPIPRAVMRLLEDLTLSBEFLAKTIKSGDTAVY DVDDDKVTVLIGEKLEIVAN"

3316..3388

/gene="trnF(gaa)"

3316..3388

/gene="trnF(gaa)"

/product="tRNA-Phe"

/note="codon recognized: UUC"

/anticodon=(pos:3349..3351,aa:Phe)

/complement(3402..5201)

/gene="dnaB"

/complement(3402..5201)

/gene="dnaB"

/codon_start=1

/trans_except=(pos:complement(5199..5201),aa:Met)

```
/product="replication helicase subunit"
/protein_id="AAC35596.1"
/db_xref="GI:3602935"
/translation="MKVTKFYNLKAERHLLIIQLIVKDFLQVTKIDIVTFYQSHQ
ILVRSINILHAKVILNPNLILHSELVETVNLILFNDRDFRLININDSLK
VLNDFIRRELONSCAKIVLTFNPNLSEVLELQKSLILNINYNKAFMLISLQ
LLEITLEDKKTNSSTVLTGFPLDLHVLQKSDLIILAGSPMKAFMLISLVN
VADIOSPPIVLEMSKOLIVLISNENIARSLRBNISISEMILNRMATLIS
NUNIVDDENDNVLDTQSKASLQOTVYGLIVDIQLQVKEKSSQGRLELVI
TRYILKLNKDNPLVLSQLSRVNERRLNKRPILSLKSGCLSTATNPYILKSN
VNFSLTSLKVNHHYILSATLELLIPQYNTIYTVSLIKELQIGKVLVLEDFIS
VIVSTVEHYVLTNGWKRIELFVDDVLADIDQVLLYNNTEVDLFSSNIVFSSVIL
ICNRIINYDFMIPKTNPFYNALVHNSLEODADVLLFLYRDSYQHNKOSINDMC
EVIIVAKHGTIGMNLINFNNTVSPNMLIKES"
/misc_feature
  /note="signal element intein; encodes putative intein
  spliced out of the replication helicase subunit"
  3594..4073
gene
  /gene="erm(cca)"
  5361..5433
tRNA
  /gene="erm(cca)"
  5361..5433
  /product="tRNA-Tyr"
  /note="codon recognized: UGC"
  /anticodon=(pos:5394..5396,aa:Tyr)
  5588..6013
  /gene="tp111"
  5588..6013
  /gene="tp111"
  5588..6013
CDS
  /gene="tp111"

Query Match      28.2%; Score 159.6; DB 8; Length 121524;
Best Local Similarity 60.3%; Pred. No. 8e-34;
Matches 298; Conservative 0; Mismatches 194; Indels 2; Gaps 2;
```

```
DEFINITION
  Streptomyces cinnamonensis acetohydroxy acid synthase small subunit
  (ilvN) gene, complete cds.
ACCESSION
  AF175526
  AF175526.1 GI:5733115
VERSION
  AF175526.1
KEYWORDS
  Streptomyces cinnamonensis
SOURCE
  Streptomyces cinnamonensis
  Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
  Streptomycetaceae; Streptomycetaceae; Streptomyces.
REFERENCE
  1 (bases 1 to 593)
  Kopecky,J., Janata,J., Pospisil,S., Felberg,J. and Spizek,J.
  Mutations in two distinct regions of acetolactate synthase
  regulatory subunit from Streptomyces cinnamonensis result in the
  lack of sensitivity to end-product inhibition
  Biochem. Biophys. Res. Commun. 266 (1), 162-166 (1999)
JOURNAL
  MEDLINE
  PUBMED
  10581183
REFERENCE
  2 (bases 1 to 593)
  Kopecky,J., Janata,J., Pospisil,S., Felberg,J. and Spizek,J.
  Direct Submission
  Submitted (05-AUG-1999) Genetics and Physiology of Actinomycetes,
  Institute of Microbiology, Videnska 1083, Praha 4 CZ-14220, Czech
  Republic
FEATURES
  source
    location/Qualifiers
    1..593
    /organism="Streptomyces cinnamonensis"
    /mol_type="genomic DNA"
    /strain="C-100-5"
    /db_xref="taxon:1900"
    30..557
    /gene="ilvN"
    30..557
    /gene="ilvN"
    /codon_start=1
    /transl_table=1
    /product="acetohydroxy acid synthase small subunit"
    /protein_id="AAD4932.1"
    /db_xref="GI:5733116"
    /translation="MSTHTLSLVENKRGVLTALFRRGFIDSLAVTEHPD
    IRRITIVNVEDLPBOYTKQNKLVNKLKVELEPAGAAVARELVAVRDNERSQ
    IVEIVQLEPAKTVDSPEAVTIEATGSSDKLEPMLKMLEPFEIKELVSGTILGRGS
    RSITRSLNALDRSA"
ORIGIN
  Query Match      27.0%; Score 152.4; DB 1; Length 593;
  Best Local Similarity 59.6%; Pred. No. 5.7e-32;
  Matches 292; Conservative 0; Mismatches 196; Indels 2; Gaps 2;
```

Mon Jul 26 12:13:47 2004

us-09-893-033-17.rge

Page 25

Db 394 TCACCATCGAGGCGACCCGGACGACGAGCGACAACTGAGAGCCATGCTCAAGATGCTGGAGC 453

OY 465 AAGTTGGCAATTAAAGAGGTGGCTCGAACGGGCAAAATTGCTTTGGTGGCGAAATCCGGCG 524

Db 454 CCTTGGGCATCAAGAGAGCTGTGCGATCGCGGACCATCGCATCGGCGCGGCTCCGGCT 513

OY 525 TCATATAGCGA 534

Db 514 CCATCAGCGA 523

Search completed: July 25, 2004, 10:17:21
Job time : 2273.95 secs

